Clinical judgements by nurses : decision strategies and nurses appraisal of patient affect.

Openshaw, Sheila

Clinical Judgements by Nurses: Decision Strategies and Nurses' Appraisal of Patient Affect

by

Sheila Openshaw

Thesis submitted for the degree of Ph.D

University of London

Chelsea College

August 1984

Thesis Contents

Acknowledgements

I should like to express gratitude to Dr Jenifer Wilson-Barnett. This research would not have been possible without her help and guidance. I am most grateful for her supervision.

I should also like to thank Professor Andrew Mat hews for his time spent giving advice and suggestions. My typist Janis House, has also been very generous and patient.

I would like to acknowledge all the help I received from the staff and patients who were involved in this research.

Lastly, I would like to acknowledge the Department of Health and Social Security and the Wellington Foundation who funded this research.

Summary

The work reported within this thesis was concerned with two types of clinical judgements that are made by nurses.

The first part of the thesis examined the comparative utility of adopting a particular search strategy when diagnosing the presence or absence of a post-operative complication. The strategy was 'Successive Scanning' which requires a reduction of possible hypotheses by the initial selection of appropriate, general information. This search strategy was examined in a pair of experiments that employed a recognition paradigm, thus separating the effects of the strategy employed from those due to limits imposed by memory processes. It was demonstrated that the people who asked the general questions that they posed earlier in the task asked fewer questions overall. Any association with increased accuracy was equivocal and this was discussed in relation to the methodology employed and future work.

The second part of the thesis was a descriptive, correlational study of nurses' perceptions of patient affect. A judgement that would influence the quality of interpersonal care that a nurse can give to patients. The methodology employed involved the rating by 80 nurses and 140 patients of a Mood Adjective Check List that contained 5 moods ie. vigour, fatigue, anxiety, depression and hostility. The patients rated the check list to describe how they felt that day and the nurses rated it as they considered that their patients had felt. The correlations obtained indicated that nurses, on the whole, are more able to describe the physical than emotional responses of their patients, especially tiredness, and this skill increases with experience. Anxiety was considered overestimated by all the nurses. The role of stereotype use and personal identification with the patient were examined in relation to these judgements.

The conclusions that emerged from the first half of the study were suggestions for further work with design modifications that would allow clarification of an association between adopting a particular strategy and accurate performance. The second part of the study provided an initial overall description of the nurses' judgements of 5 different patient moods. Suggestions were made for a more detailed analysis of patient affect. The implications of the work were directed towards nurse education.

Introduction

A considerable body of research has accumulated documenting the fallibility of human judgement in clinical and non-clinical settings (Einhorn, 1972: Goldberg, 1968: Slovic, Fischoff & Lichtenstein, 1977). There has been controversy about the accuracy of predictive clinical judgements made by doctors, psychiatrists and clinical psychologists. Evidence has been cited suggesting that the clinical predictions made by these professionals under the inevitable conditions of uncertainty were made as well, if not better, by the application of a simple statistical formula (Meehl, 1954). The implication was that the sign and symptom information upon which a clinician relies was handled better by a statistical model. Higgins (1980), discussing decision making in general practice, made the distinction between two processes that should be involved in a medical diagnosis. The collection of sign and symptom information and the doctor/patient contact. He suggests that the latter area is neglected in much medical training and the evidence from psychology concentrates mainly on the former.

The issues surrounding clinical judgements are important to nurses. A necessary pre-requisite for a nurse effectively caring for a patient is that she accurately perceives that patient's current need. This has obvious implications for a 'nursing process' or problem solving approach to patient care (McGilloway, 1980). Identification of a patient problem is dependent upon a decision by the nurse about the current physical and emotional state of the patient in her/his care. Further, this would help to clarify a definition of nursing 'care'. If nursing care is seen as starting with the identification of a patient problem, it is necessary to ensure that nurses are equipped to do this.

Of added importance is the trend in North America to extend the role of the nurse/nurse practitioner. This is, in part, a response to economic pressures reflected by the rising cost of health care and a non-uniform distribution of family doctors. The result for the nursing profession is that nurses are making decisions that were previously the province of medicine. With a similar economic climate in this country and a concomitant desire on the part of some nurses to expand their role (Bowling, 1981), a similar trend could emerge in Britain. This makes an examination of clinical judgement and decision making of particular interest to nurses. But nursing decisions are very varied. They involve the manipulation of sign and symptom information leading to diagnosis of a physical problem. They also involve evaluation of their patients' psychological well being.

Higgins (1980) had noted the lack of emphasis on the doctor/patient communication component of clinical judgement. For these reasons the work reported within this thesis focused on both types of judgement, but separately rather than simultaneously. This was considered to present a more realistic approach to the study of clinical judgements made by nurses. The first half of this thesis reports two experiments examining a strategy employed by nurses when starting to seek information that leads to the making of a physical diagnosis. The literature background is that of decision theory and information processing. The research focused on the consequences for judgement of adopting a specific strategy to initiate the process. The second half of the thesis reports a shift of emphasis from physical diagnosis to that of the judgement of affective well being. The focus is concerned more with the accuracy of nurses' perception of their patients' feelings than the strategy adopted in forming that perception

CHAPTER I

Review of the Literature for the Experiments

Diagnosis or clinical judgements are predictions or decisions made about uncertain events. There can be no such things as error-free diagnosis because a cue (symptom) and the criterion (disease state) with which it is associated are not functionally related. That is, any symptom or sign does not have a one-to-one relationship with a particular disease or patient condition. It is typically indicative of more than one disease state. It is the configuration of the sign and symptom cluster and the relative weights of the individual cues that suggest a diagnostic category (Einhorn, 1972;1974).

The areas of clinical judgement and Behavioural Decision Theory encompass an extensive body of work and a complete review is beyond the scope of this study. Kleiter, Gashowetz & Huber (1976) cited in excess of two thousand individual references at that time. It is intended to select a relevant sample of the various work in the area as a whole, dividing the literature review into discrete sections. The objective of the review is to evaluate the theoretical frameworks that have been proposed to describe decision making and to consider decision aids currently in use with the aim of isolating a method that could be applied to facilitate a nursing decision.

In order to consider experimental evidence concerned with decision making and judgement it is necessary to have a conceptual framework within which to operate. Hogarth & Makridakis (1981) proposed the framework illustrated in Figure 1.

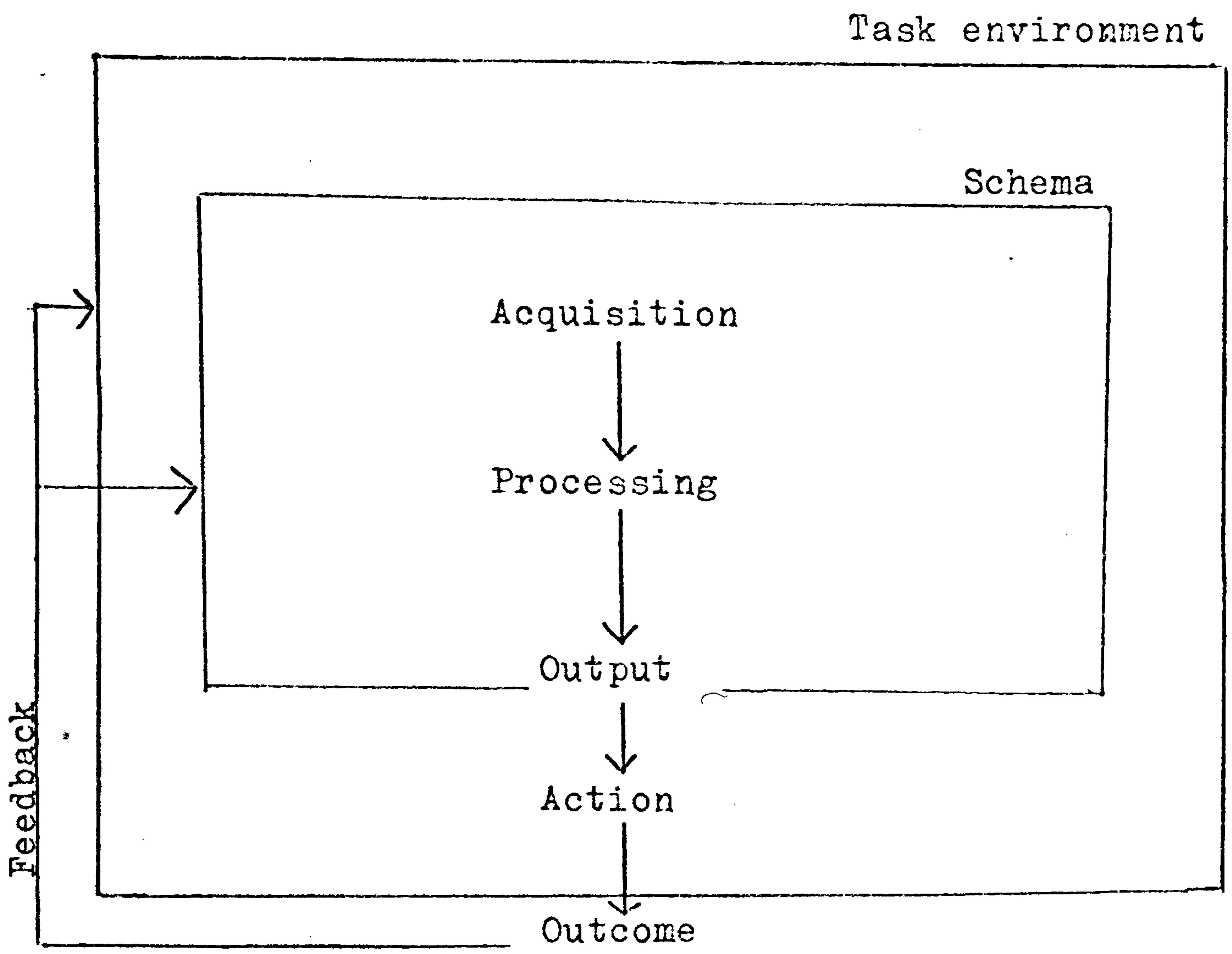


Figure 1

Model of Judgement from Hogarth & Makridakis (1981)

The operations that are involved in making a decision or judgement can be broken down into different stages. Firstly there is the acquisition of information, followed by the processing of the information that has been accessed, and then output. Output implies taking an action based on that decision. This action and external factors present in the environment result in an outcome. The knowledge of that outcome feeds back into the decision maker's schema or mental representation and can also affect the task environment. At the same time the schema influences the perception of the task environment and its complexity, the problem identification, the tasks involved, the type of actions required and the objectives of those actions.

This general model provides a framework within which to discuss the processes involved in making decisions. The evidence will be reviewed with reference to this model, although not in the logical sequence, rather, an historical approach will be adopted. The type and range of judgements studied will be discussed initially, followed by the models adopted that eventually resulted in theories of decision making. The decision environment and cognitive factors will be discussed, followed by theories of decision making. Work specifically relating to nursing and medical decisions and decision aids will be considered, although the clinical relevance will be referred to throughout. Then the whole review will be related to the stages of the general model. The final section will be concerned with an appraisal of aiding decision making, and the rationale for the experimental work.

1.1

Judgements Studied

The area divides broadly into the study of clinical judgement ranging over a variety of different disciplines and the study of judgements adopting a gambling or "risky choice" paradigm. The whole area of judgement/decision making will be considered in relation to an identification of underlying principles and unifying theory. The types of judgements studied have included forecasts made by economists and business managers, the selection of personnel and post-graduate students, as well as clinical judgements or diagnoses made by clinicians.

However, all the work shares the common factor that it is concerned with "clinical" judgements or predictions rather than objective projections. All the decisions rely on information that can be incomplete, is often subjective and is invariably probabilistic. As was stated earlier, by probabilistic is meant that there is no functional relationship between a criterion and the cues that relate to that criterion. In terms of a nursing decision, what the nurse has to decide is the liklihood that a particular sign or symptom predicts a particular patient state.

Abelson (1976), proposing a theory of decision making, noted that remarkably little was known about how people make decisions. The study of decision making had been dominated by the model of "rational man" in economic theory. This approach has been less than adequate, as will emerge. But, it has provided useful insights into what people cannot do and insights into some of the "biases" that operate as judgemental heuristics or strategies in certain circumstances and form the basis of recent theory. It is arguable that a normative prescriptive approach adopted by decision theorists will ever be adapted to describe the cognitive processes that take place when people make decisions. But normative and statistical models have prescribed a criterion of objective performance, that has been assumed to be optimal, against which a natural decision can be compared. Inferences about decision processes were made when the human decision failed to correspond to that prescribed by the model. The approach is much the same as that adopted in the study of human memory, where failure of recall has been suggestive of possible mechanisms.

The danger of this approach for theory is when prescribed or potential competence in a situation is confused with naive expectations of the descriptive power of a normative model, ie. how the task should be performed by the person. But the usefulness of normative procedures in prescribing and aiding certain decisions, ie. medical decisions, is very real, as will be discussed. But what does characterise much of the work in this area is a lack of performance invariance. This has made the identification of a small number of underlying principles, that can describe observed behaviour complicated (Payne, 1982). The following sections will expand on this theme.

1.2

Early Models Adopted

During the 1950's there were many studies comparing judgements/decisions made by actuarial techniques with those made by clinical techniques. Meehl (1954) reviewed all the studies prior to that date and arrived at the conclusion that no study gave a clear superiority to clinical techniques. Actuarial predictions were consistently as good, if not superior. The ability of people to combine information from different sources when making a prediction was no better than applying a simple statistical formula to that information. As often as not, it was worse.

Holt (1958) presented data describing the predictions of success of individual doctors engaged in psychiatric training. He found that, in this instance, clinical procedures produced superior predictions. This was an isolated instance of superior clinical performance at that time and posed the question of the types of decisions best suited to clinical techniques. Richards (1963), in an attempt to answer this question, suggested that statistical techniques are maximally sensitive to differences between classes of people or items. They have an accurate stereotype of the predicted behaviour of the members of a class of items. Whereas, predictions based on clinical insights are best suited to situations where these predictions are concerned with differences between members of the same class, predictions about people who belong to the same category eg. have the same diagnosis. However, many clinical decisions are concerned with distinguishing between categories, not projecting differences in behaviour by different members of the same patient group.

The two inter-related theoretical aspects of work will be reiterated. There is Normative Theory, which is concerned with prescribing a course of action which adheres to expectations of decision competence ie. reflects the beliefs of the decision maker. The second type of theory, Descriptive Theory, is concerned with how people perform the task (Slovic, Fischoff & Lichtenstein, 1977). Slovic et al. suggested that the early descriptive studies consisted mainly of rather superficial comparisons of actual behaviour and normative models.

The considerable effort that has been devoted to studying how people perceive, process and evaluate the probabilities of uncertain events is a reflection of the assumed importance of probabilistic reasoning to decision making. The early research of Peterson & Beach (1967) on 'intuitive statistics' lead to the optimistic conclusion of a rational, if 'conservative', decision maker. This work was followed by model-based paradigms, that is normative models in descriptive research. The suggestion was to begin with a normative model and adjust its form or parameters to produce a descriptive model (Barclay, Beach & Braithwaite, 1971). The normative model was to be adjusted to reflect the processes underlying performance (Fig. 1). The study of 'conservatism' demonstrated that in certain situations people have a tendency to produce posterior probabilities that are nearer to the prior probabilities than those specified by a normative model when they are integrating probabilistic information. They acted with more caution than the statistical model would prescribe, a rational if conservative approach.

But, the question was whether or not they would adopt the same strategy regardless of differences between experimental tasks. Exceptions were soon elicited to question the generality of this finding and suggested that the 'conservatism' demonstrated by subjects was a function of the experimental task - a highly simplified task.

Real-life problems often have several stages, each stage relying on information which is itself an inference from unreliable observations or reports (Slovic et al. 1977). An example given by Slovic et al. is the physician who uses the condition of the patient's lungs as a cue for diagnosis. She/he must first infer that condition from unreliable data, eg. the sound of a thumped chest.

Descriptive studies of multi-stage inference found that people consistently demonstrated posterior probabilities that were more extreme than those prescribed by the normative model (Funaro, 1975, Peterson, 1973, Steiger & Gettys, 1972). This is in direct contradiction to the 'conservatism' findings. The extremity of response has been suggested to occur because of the use of a simple, but inappropriate 'best guess' strategy which is insensitive to considerations of data unreliability (Gettys, Kelly & Peterson, 1973; Snapper & Fryback, 1971). Subjects did not adopt and use the principles of normative theory when making intuitive decisions/judgements. But whether the strategy was that of a 'best guess' can be questioned.

Thus far it is impossible to detect how people do use probabilistic information in terms of the general model (Fig. 1), processing is inferred from output where acquisition is controlled by the experiment design. This early work can say little about how information is acquired or processed by a natural decision maker. The following section reviews further evidence of statistical naivety that was suggested to be characteristic of natural decision making. But, this work was among the first to provide examples of specific biases that appeared to serve as heuristics or decision strategies.

1.3

Heuristics and Biases

Tversky & Kahneman (1971;1973) & Kahneman & Tversky (1973), in a series of studies showed that people did not attend to or failed to learn about the statistical realities that are assumed to be present in the environment when making predictive decisions under uncertain conditions. Prior probabilities, even when supplied, tended to be ignored at worst and neglected at best. Predictions tended to be too extreme and were not properly regressive. They tended to be insensitive to the reliability of the data as a source of information. Normative theory prescribes that predictions should be more regressive, that is, less extreme with increased data unreliability. Further, people demonstrated an unjustified confidence in their judgement as a function of consistent or extreme information.

The subjects in these studies demonstrated a limited number of "biases" that appeared to serve as judgemental heuristics. Judgemental heuristics being simple strategies that were adopted by subjects when performing a decision task. The nature of the heuristic/strategy used was determined by the nature of the task, and the use of a particular heuristic being inferred from performance. In terms of the general model (Fig. 1) they may provide a suggestion as to how information is acquired but are inferred at the response stage. The more recent descriptive studies, rather than simply comparing behaviour with normative models, have attempted to uncover how the underlying decision process is affected by both the limitations of the decision maker and the demands of the task. It was the demonstration of judgement heuristics that gave impetus to this change.

A 'representativeness' heuristic or strategy is probably most important when people are deciding the category or group to which a person belongs. In the study that described 'representativeness' subjects decided whether fictional characters were either engineers or lawyers on the basis of their apparent similarity to a social stereotype of an engineer and lawyer (Kahneman & Tversky, 1973). 'Representativeness' was said to be assessed by an examination of the essential features of, for example, two items X and Y. In order to gauge the probability that item Y belongs to class X, the degree of similarity between the two items is assessed. The degree of similarity is the determining factor. When Y is very similar to X, then its probability of belonging to class X is judged to be high. Kahneman & Tversky found that a short description depicting the social stereotype of an engineer would lead people to predict that the person described was an engineer. This was despite contradictory base-rate information stating the number of engineers in the population that was being evaluated.

In other situations an 'availability' heuristic appeared to be used when making predictive decisions. An event, such as a particular disease, is judged as being likely or having a high incidence if it is easy to recall relevant instances from memory. This is often a valid cue because frequent events are often easier to recall than those that are less frequent. But availability can also be influenced by other factors that are unrelated to actual frequency, eg. the emotional saliency of a cue, the recency of previous experience or its discriminability. An example in the current context of decision making is one concerned with media coverage of killer diseases. Those diseases, eg. lung cancer and all accidents, that received more coverage by the media were rated as killing more people (have a higher incidence) than those diseases that received little media coverage, eg. stomach cancer and strokes (Slovic, Fischoff & Lichtenstein, 1976).

A third heuristic described was 'anchoring and adjustment'. A natural starting point, or an anchor is used as a first approximation to the judgement, for example after having received one or two items of information such as knowledge of a couple of symptoms. The anchor is then adjusted to accommodate the implications of additional information such as a physical examination. The anchor can be imprecise and adjustment insufficient. An earlier study had demonstrated this effect (Hammond, Kelly, Schneider & Vancini, 1967). Nurse subjects revised their judgements about the probability of a fictional patient having a particular condition about one-third as much as a normative model would prescribe when presented with a fresh piece of information.

These findings have been obtained in numerous similar studies (Wise & Mockovat, 1973; Bar-Hillel, 1973: 1974; Hammerton, 1973; Lyon & Slovic, 1976; Nisbett & Borgida, 1975). However, there have been criticisms of the work in terms of artificiality, the relevance of normative theory and post-hoc explanations of the data.

Cohen (1979) suggested that Kahneman & Tversky assume that the human mind has only one legitimate framework within which to reason about uncertain events ie. probability theory. He stated that to accuse someone of a computational error within a logical or mathmatical system, you need first to be sure that you have correctly interpreted what system is being used. But, in order to study decision making it is necessary to have an objective criterion of competence. The simple normative model, Baye's Theorem, serves to set a criterion of potential competence against which a natural decision can be compared. The problem arises with an assumption of optimal performance and optimal models. At the end of the day what are considered to be optimal outcomes are evaluated by judgement. Outcomes as a result of a process that is considered to be rational and desirable, and the axioms are thought to be reasonable (Einhorn & Hogarth, 1981). It is with this qualification that apparent deviation from a normative prescription can be interpreted. A normative approach, despite descriptive shortcomings, has stimulated considerable work but whether it reflects the natural environment is questionable. Kahneman & Tversky (1982a) identified a need for a positive account that explains errors via heuristics and a negative account explaining why the correct rule has not been learned. It appears to have been the emphasis on the notion of 'error' or 'bias' in judgement that has been controversial. These points will be returned to at a later stage.

Olson (1976) suggested that a fundamental problem with the 'representativeness' heuristic is in inferring use from a post-hoc analysis of the data. It is necessary to have a prior specification of the essential characteristics of the salient features against which the judge will assess 'representativeness'. This criticism does apply to the numerical prediction studies, but is less cogent with respect to category prediction. In this case, the 'representativeness' of a social category was an independent variable that was systematically varied, although those characteristics that were essential have not been precisely identified. The strategy employed by people in a condition of an uninformative description, who still ignore base-rates cannot be interpreted from this evidence. Rather, the mode of presentation of base-rate information is worthy of consideration and this will be discussed in the next section.

Overall, the strategies or heuristics that subjects employed can be inferred from performance as an initial step in examining how natural decisions are made. It is interesting to examine why people fail to learn about or are unable to use statistical realities, if this is the case, and apply them to prediction. The evidence reviewed thus far suggests that the heuristic or strategy selected, the way it is employed and the "accuracy" of judgement are all very problem specific. Heuristics are inadequate as a general theory of judgement because of the difficulty of knowing which one will be applied in any particular situation. However, whether a general theory will predict the strategy employed in all situations is open to debate. The next section considers why people fail to learn the conditional probabilities of the decision environment, and questions further the use of optimal models.

1.4

The Task Environment

The task environment in which nurses and doctors typically function will be discussed in relation to the persistence of an "illusion of validity" that is commonly held about judgement accuracy. Self-confidence in judgements made by a group of clinicians has been found to increase as a function of the amount of information available to them, but without any corresponding increase in judgemental accuracy (Oskamp, 1965). People have been found to show most confidence when information is consistent and/or extreme. But the presence of extreme information is a situation of maximum regression, when maximum caution and least confidence should be exhibited (Kahneman & Tversky, 1973).

Castallan (1977) emphasised the importance of understanding the environmental characteristics in which behavour occurs. The decision making environment is very complex, with many factors operating. Estes, (1976) stated two general conditions for an accurate estimation of probabilities:

1. Alternative events involved in a situation must have equal opportunities of occurence, and

2. the learner must attend to and encode occurences of all the alternative events with equal uniformity and efficiency"

However, these preconditions are violated when a judgement leads to an action. The choice of one alternative excludes others, therefore, outcomes that are contingent on actions taken and not taken will typically have unequal opportunities of occurence. Further to this the independent variables in the environment must be disentangled by cognitive activity, and therefore be subject to effects such as memory loss, information overload, recency and primacy effects (Hammond, 1978).

Einhorn & Hogarth (1978) examined the decision environment comprehensively where the basic tasks considered involved judgements made for choosing between actions. They observed that most studies simply correlated judgements with criteria and neglected choice between actions. A model for learning and maintaining confidence in judgement was the result and the role of positive (successful decisions) and negative (unsuccessful decisions) outcome feedback were considered. In the real world, decisions are made for the purpose of choosing between actions. The outcome feedback, or information about the accuracy and appropriateness of the decision, only being available after the action has been taken and is often the only available source of information.

Evidence was cited that suggested that people judge the strength of a relationship by frequency information rather than probabilities (Jenkins & Ward, 1965; Smedslund, 1963;1966; Ward & Jenkins, 1965). It has also been shown that people have difficulty using disconfirmatory evidence (Wason, 1960; 1966; 1968; 1969). Einhorn & Hogarth (1978) suggested that the principle difficulty lay with the structure of judgemental tasks in the natural environment as this determines the conditions in which inferential learning can occur. Where a nurse selects patients for a treatment or people for nurse training, what is unknown is the response of people for treatment or training who were not selected.

The model developed relating positive hits (correct choice of action) and false choices of action (false positives) to confidence in judgement made two basic assumptions. Firstly, evidence about outcomes that are contingent on actions that are not taken is frequently missing, or if the outcomes are available attention is not paid to them. Secondly, action/outcome combinations are coded as frequencies rather than probabilities. The model considered two possible actions denoted A and B, where χ was an overall evaluate judgement and χ c was a cut-off point. Therefore:-

$$\text{if } \chi \geq \chi c \text{ choose action A}$$
$$\text{if } \chi < \chi c \text{ choose action B}$$

In order to compare decision χ to a standard, the existence of a criterion Y was assumed to serve as a basis for evaluating the accuracy of judgement ie. outcome feedback with cut-off point Y c, therefore $Y \geq Y$ c and $Y < Y$ c. The two types of error would be:-

false positives $(Y < Yc / \chi \geq \chi c)$ and
false negatives $(Y \geq Yc / \chi < \chi c)$

The probability of observing a successful outcome was stated to depend on:-

1. Y c, the base rate
2. χ c, the selection ratio
3. $\rho\chi Y$, "true" judgemental ability ie. the correlation that would occur in the absence of treatment effects
4. t, the size of the treatment effects

These four factors were systematically varied in a simulation experiment and the observable positive hit rate was shown to be maximally affected when judgement was highly selective (low selection ratio) and when the base rate and selection ratio were close to each other. Figure 2 illustrates the effect of treatment and the assumptions of the model upon observable judgement.

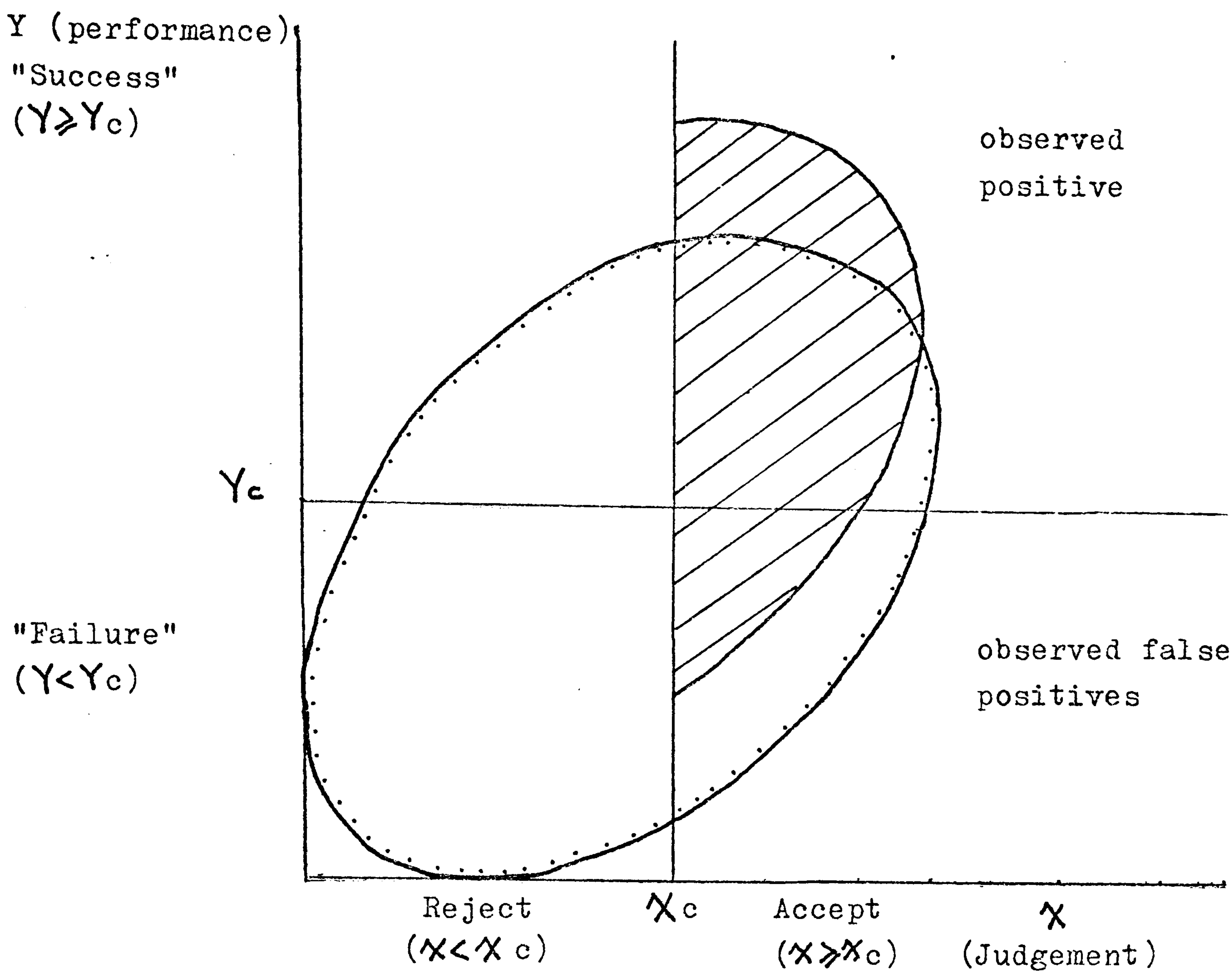


Fig. 1. Effects of treatment on the observed positive hit rate (from Einhorn & Hogarth, 1978)

The dotted elipse represents the "true" relationship between judgements and outcomes and the shaded portion indicates only the outcomes that can be observed. Where treatment is weighted to reflect its relative size all actions given A are increased by a constant amount, so the observed number of positive hits is greater than would be the case in the absence of treatment. The feedback to the decision maker is contaminated, in that the number of positive hits is inflated with a concomitant obscuring of false negatives, those who would have responded had they been selected. The rationale behind the methodology of chemotherapeutic research is to clarify the observation of outcomes. Where two drugs are being compared with each other and (ideally) a placebo, a substantial group of patients with a comparable disease status are randomly ascribed to one of the three conditions.

The main conclusions from Einhorn & Hogarth (1978) can be summarised. Outcomes that are contingent on actions taken and not taken will, typically, have unequal opportunities of occurence. If there is equality, disconfirming evidence may be ignored. If symptom X implies disease Y, Y is unlikely to be considered in the absence of X. The simulation results indicated high observed positive rates in many situations. Also, if information in the natural environment is acquired sequentially ie. is processed item by item, memory for disconfirming instances needs to be aided.

Judgements and actions are taken in particular task environments. Factors, such as base-rates, selection ratios, treatment effects, uncertainty of the task, serial vs, simultaneous presentation of information, completeness of judgement/action combinations can all vary between different environments. The combination of judgements, actions and environments produce outcomes. If awareness of environmental variables and their effects is lacking, outcome feedback will be ineffective. There is work to show that this is the case (Castallan, 1977; Slovic & Lichtenstein, 1971). Also, in the absence of adequate control or understanding of environmental factors, inference regarding causal relationships between judgements, actions and outcomes is problematic. Therefore, if in many situations people are unable to observe and learn the real distribution of outcomes and have insufficient evidence to refute the validity of their judgements, an illusion of validity will persist and will be difficult to change. People do not have the opportunity to assess probabilities.

Einhorn & Hogarth (1981) questioned the prevelance of normative models in the study of decision making and their absence in other branches of science. As discussed earlier, they isolated a need for a criterion of optimal performance. An optimal criterion where behaviour is assumed to be purposeful and goal-directed, the objective being to attain those goals with maximum benefit and minimum cost. In other words an assumption of functionality. They defined optimality as:-

> "...., decisions or judgements that maximise or minimise some explicit and measurable criterion (eg. profits, errors, time) conditional on certain environmental assumptions and specified time horizon".

The emphasis was on the conditional nature of optimality. Simon (1979) observing the complexity of the environment had suggested that it is necessary either to construct optimal models by making simplifying environmental assumptions, or maintain greater environmental realism through heuristic models.

Einhorn & Hogarth (1981) observed a lack of consideration of the conditional nature of optimal models. They took the frequent finding that people are insufficiently regressive in their judgements and suggested that the situation of a changing process, as opposed to stability, had been ignored. They suggested that with a changing process regressive predictions are suboptimal. The example given was, where large financial losses suffered by a company are attributed to be due to poor management and worsening conditions, a realistic forecast would anticipate even more extreme losses. The process of change is mediated by the decision makers' attribution of causality. A statistical model would assume that the recent losses were generated by a stable process and would predict that profits would regress to their mean level.

Further assumptions surrounding a normative model's prediction of optimal performance were criticised in this review. Much of the work assumed conditional independence, perfectly reliable data and well-defined sample spaces that may not be representative of the information in the natural environment. Other work has questioned an interpretation from the studies employing statistical models that suggested that people inevitably ignore base rates (Christensen-Szalanski & Beach, 1982). The abstract nature of a base rate summary statistic used in many of these studies may not have been salient enough for the subjects to recognise the relevant population that defined that base rate. In an experiment where pre-training was given about base rate information, subjects learned and used that information. Also, physicians' diagnoses of the presence/absence of pneumonia suggested that information about the base rate prevalence of pneumonia was being used (Christensen-Szalanski & Bushyhead, 1981). Further the mode of presentation of information about risk influenced the perceived importance of that risk when subjects estimated the importance of five diseases as causes of death (Harding, Eiser & Kristiansen, 1982).

Overall, environmental factors have been largely ignored when predicting optimal performance. The outcome feedback that people observe often prevents them from learning about "true" relationships and those "true" relationships only hold under specific conditions that may not be reflected in the natural environment. These criticisms were directed largely at assumptions of competence predicted by statistical models. A normative cost/benefit theoretical approach attempts to make allowances for the environmental limitations imposed upon the cognition of the decision maker (Beach & Mitchell, 1978) This will be discussed further in a later section, the issue for a medical decision is that base rate information can be neglected and decision making can be aided by the use of a simple formula. The next section discusses a broader range of work that explores further cognitive factors that influence the way in which information is acquired and processed.

1.5

Information Processing

As was stated earlier, the need for a dual account of the findings of experiments employing normative models has been recognised (Kahneman & Tversky, 1982a). They suggested a need for a positive account that explains errors via heuristics and a negative account explaining why the "correct" rule has not been learned. The preceding discussion was concerned with the context of the decision environment and there is a large amount of work that has studied how people process and learn to use probabilistic cues when making judgements about events (for a short review Castallan, 1977). The early work adopted a binary-value single cue task, ie. the cue was valued at either 0 or 1 (Atkinson, Bogartz & Turner, 1959; Burke & Estes, 1957; Atkinson, 1961; Summers, 1968). But this simplified methodology could only touch on the problem. Other work was concerned with concept learning under conditions where the feedback that was given was either informative or misinformative.

A combination of the early paradigms has resulted in studies that involve people in multi-cue probability learning (MCPL) tasks. These tasks required subjects to integrate information from two or more stimulus dimensions, where the validity of a cue predicting an event must be inferred. The results from the studies found that subjects attended to both dimensions in a task where two were used and further, they attended to both even when one was less relevant (Castallan & Edgell, 1973). This interferred with performance most when the relevant cue was of low validity (Castallan, 1973). The overall findings from the MCPL studies demonstrated that people can learn to process information about cue dimensions and configurations of cues when it is appropriate to do so. But, in probabilistic judgement tasks, subjects did not adopt optimal response strategies. They did not learn to ignore irrelevant information, even when they were told that it was irrelevant (Castallan, 1977). Knowledge about the structure of the task did not necessarily lead to learning.

Models and theories of decision making will be reviewed in the following section, but Estes (1976) noted that no single model of decision-making could fully account for all the probability learning data to that date. Although different models did account for and predict portions of the total data of the vast number of experiments in this area, Estes pointed out that in all cases, what is learned by the judges has to be inferred from performance in an experimental situation and concluded that the likelihood of one general model accounting for all the findings was low. In order to try and understand why different models were required to deal with different situations, Estes took a step back and considered the influence of memory on performance. He considered how and what people learn about probabilities.

Tulving (1972) postulated a distinction between two components of memory, an episodic store and a semantic store. These stores being seen as conceptually, as opposed to physically, distinct. The episodic store is concerned with the recording of instances of events whereas the semantic store is concerned with meaning, or, in the current case, relative frequencies. It was apparent to Estes that different models were concerned with different components. But the typical probability paradigm represents a mixture of contributions from episodic and semantic memory. When the effects of the two different components were experimentally separated, subjects were found to be very efficient at storing information about probabilities as relative frequencies. They appeared to scale the frequencies of events in memory. But stimulus per se did not influence the choice of probability. Estes suggested that subjects categorised the events involved in a task and then learned the relative frequencies within classes. But they did not always carry the process of categorisation far enough, possible because of limited memory performance.

Sniezek (1980) attempted to distinguish between the ability to acquire information about probabilistic events and the ability to use that information. In an experiment where subjects had to learn the relationship between cues and the criteria that were predicted by those cues, they were also required to retrieve information from memory about the probabilistic relationship between the cues and criteria. Performance demonstrated that subjects could ascertain the probabilistic information, however, this ability was negatively affected if they were asked to recall specific past events. The cognitive load required to simultaneously learn and apply probabilistic relationships, plus make a prediction, was too much for these subjects. The materials involved in the experiment were all numerical and the implied application of the probabilistic information required arithmetical calculation and manipulation which made this a very difficult task to do.

The implication is that the "expert" in a clinical judgement situation, in this case a doctor or nurse, can successfully extract information about signs and symptoms and the patient condition that they signify. It is the combination process that could be problematic. Work looking at the role of the "expert" in a decision task made the distinction between data collection and the method of combination of that data (Einhorn, 1972). Einhorn's study was an intensive investigation of three pathologists who each viewed the slides of 193 patients who had died as a result of Hodgkin's disease. Their task was to predict survival time in months on the basis of nine histological signs that the subjects had previously specified. The patients had all had histological investigations on entry to hospital, and their survival time was known. The nine components used did correlate with actual survival time. However, the clinicians overall global judgement of severity did not show the anticipated negative correlation with actual survival time, rather the correlations were random.

The thrust of Einhorn's work was to consider the role of machines in aiding the task of diagnosticians. The suggestion being that the expert defines the relevant cues that need to be combined, and the combination is done mechanically. Clarke (1982) suggested decision analysis for doctors as a supplement to the acquisition of scientific knowledge in the education of surgical residents.

Taking the mechanisms underlying information processing a stage further, it is now intended to consider characteristic encoding biases that occur as a result of the categorisation process. In terms of the general model the concern is with selective acquisition and processing. Grosz & Grossman (1968) studied the response of five pyschiatrists and found that they each had an idiosynchriatic way of handling clinical data. They found a tendency for the individual clinicians to show a consistent overall bias in their judgements of routine psychiatric examinations, a bias that was independent of the data. This small study gave an indication of an effect worthy of more systematic study. Hammond (1970) investigated the poor reliability of judgements made by clinical psychologists, attempting to identify the sources of variance that result in poor reliability. Hammond made the point that there is a difference between judgements of physical attributes, eg. size and weight, and a personality variable, eg. dependency or aggressiveness. The former are identifiable and measurable by virtue of possessing a generally agreed-upon definition and procedure for assessment, whereas, the latter are not.

As has already been discussed the context of occurrence alters expectations. In a study of perceptual processes, subjects viewing a rapidly rotating cross saw a cross or a square, dependent on what they were told to expect (Hochberg & Brooks, 1977). This suggested integration of information into a schematic map appropriate to expectancies, which in this case was probably not verbally mediated. Hammond observed a highly significant statistical relationship between cognitive structure and the concepts that the clinicians inferred from the clinical material. Gillmore & Hill (1981) found that nurses' reactions to patients with pain were affected as a result of an ambiguous diagnosis. The influence of four sociological factors, characteristics of the patient, of the clinician, the clinician's interaction with the profession and the health care system have been reviewed in relation to medical decision making (Eisenberg, 1979).

Taking this work a stage further, Arkes & Harkness (1980) liken diagnoses to schematic representations. Bransford, Barclay & Franks, (1972) found that people falsely recognise having seen related idea units as being contained within the same sentence, even when they had been presented independently. They naturally integrated related ideas into a whole. Arkes & Harkness found that the effect of reaching a diagnosis was to falsely recognise related symptoms that had not been originally presented in the judgement task. Subjects also displayed more confidence that they had seen related symptoms before than they did at having seen unrelated symptoms. However, there was a positive effect of more accurate rejection of new unrelated evidence.

Medin, Altom, Edelson & Freko (1982) in an experiment involving initial training about a fictitious disease, some symptoms were correlated with each other and others were independent. When choosing new cases, subjects tended to choose the case that preserved the correlation in preference to the case that broke it, even when the case with correlated symptoms had fewer typical symptoms. When judging which disease was present in a single case, subjects diagnoses were determined by the correlated symptoms. However, Johnson, Hassebrock, Duran & Tholler (1982) in a thinking aloud paradigm found that the treatment of ambiguous information by doctors varied as a function of experience.

It can be concluded that people do not merely encode the original symptoms. The diagnosis is a schematic representation that is a coalescence of related information. An implication is that consistent symptoms will later be thought to have been present when the diagnosis was made, even though they may not have been actually present. The work of Sulin & Dooling (1974) suggests that this effect will increase with time. They found that a judgement was better recalled than the information on which it was based. The task and environment structure plus memory organisation form a context that can suggest how the "availability" heuristic is mediated.

Pollard (1982) considered the similarity between the "availability" heuristic and a "matching" bias used to describe subjects' response in syllogistic reasoning tasks. He considered work looking at abstract and thematic syllogistic reasoning and interpreted the results in terms of the "availability" heuristic. Abstract material typically induced a response from subjects that appeared to be mediated by "matching" the available material in the stimulus presentation. The responses to thematic material appeared to be mediated by the availability of representative past experience and context. The work reviewed above studying memory organisation would predict this effect. There is a need for further experimentation regarding the conditions under which "availability" or "representativeness" would mediate response and the interactions between the two heuristics.

1.6

Descriptive Theories and Models of Decision Making

The evidence that has been reviewed has documented some descriptive shortcomings of normative models and corresponding problems of optimal performance. But the work has isolated the use of strategies that may serve as a basis for descriptive theories.

Any general theory of decision making would have to explain and predict the processes involved in the acquisition of information, processing that results in a decision and external factors that can affect both these stages (Fig. 1), and incorporate a description of the values and beliefs of the decision maker. Kahneman & Tversky (1982b) suggested that recent psychological studies of judgements made under uncertainty have treated all forms of uncertainty in terms of a single dimension of probability or belief. Berkeley & Humphrey (1982) identified seven different forms of uncertainty about outcomes that are contingent on preceding and subsequent events, the effect of extraneous events, the role of prior information, how to conceptualise worth, how to procede, how to act and the extent of a persons ability to influence change. The scope required of a general theory coupled with a precise knowledge of processes suggests a reason why there is no one theory that adequately accounts for all the findings in this area.

The work reviewed so far has, in the main, employed the statistical model Bayes' theorem as a normative model prescribing optimal performance. Other work looking at the risky choice and gambling paradigms has applied other formulations of cost-benefit principles to strategy selection that have led to a theoretical framework (Beach & Mitchell, 1978). A variety of simple strategies have been identified that have not been discussed within this review, rather the emphasis has been on those strategies relevant to the theoretical base of the study to be described.

The cost-benefit theoretical framework of Beach & Mitchell (1978) has resulted from work with model based paradigms and adjustment of parameters to meet descriptive requirements. The framework suggests that strategy selection is as a result of a compromise between the desire to make a correct decision and the desire to minimise effort. Payne (1982) reviewed studies of contingent decision behaviour and theoretical models that have been postulated to account for findings. The model suggested by Beach & Mitchell (1978) related the value of a correct decision and the costs of applying various decision strategies to eight variables. These variables pertained to the task structure and context, and the assumption was that the process selected would maximise the anticipated benefits or significance of a correct decision against the cost of using the process. The combination of cost/benefit considerations was assumed to follow an additive rule. These ideas have been further reviewed (Einhorn & Hogarth, 1981).

Payne (1982) reviewed several questions regarding which factors are most important to strategy selection, whether selection is a compensatory or non-compensatory process and the extent to which strategies are selected at the start or at multiple points during the process. It was suggested that the notion of strategy selection as a higher level decision that involves consideration of costs and benefits provided a framework for task effects and contingent processing behaviour. It also maintains an assumption of rationality on the part of the decision maker. The use of a suboptimal decision rule would be seen as the result of a rational cost-benefit approach to strategy selection (Christiansen-Szalanski, 1978; 1980). This framework is also in line with the early decision research (Simon, 1955; Bruner, Goodnow & Austin, 1956). But the meaning of cost/benefit is dependent on the task representation, thus context. A tax cut could be seen by an individual as a reduced loss from income or a welcome gain. Further, this analysis can always be applied after the event and can therefore be circular (Einhorn & Hogarth, 1981).

Tversky & Kahneman (1981) suggested that people were normally unaware of alternative frames of reference and the potential attractiveness of options. Through a simple rewording of a decision problem subjects were inferred to have been employing a gain/loss encoding of possible outcomes rather than an encoding in terms of absolute wealth. In a decision task where subjects had to decide between two alternative immunisation programmes against a lethal disease, rewording the problem from a statement of lives saved to one of lives lost produced differing results. Two sets of the two alternative programmes were used, the first set was stated in terms of the number of lives saved and the second in terms of the number lost, but the sets involved identical statements of risk. Tversky & Kahneman suggested that the change in wording resulted in a pronounced shift from risk evasion (where problem stated as lives saved) to risk taking (where stated as lives lost).

They proposed Prospect Theory, an extension of the older expected utility model, which distinguished two phases in the decision process (Tverksy & Kahneman, 1981). An initial phase where the problem is edited into a simpler representation followed by a phase of evaluation. The editing phase is seen as the primary source of context effects in decision making. The same set of options might be edited in different ways depending on the context in which it appears. Once the editing phase is completed, the basic evaluation is invariant across representations. The editing phase can be seen as processing of context information at the acquisition stage of the process (Fig. 1).

Tversky & Kahneman (1981) suggested that the editing phase is a perceptual and pre-conscious process and made an analogy with visual perception. Perception is veridical relative to the reference point of the observer ie. the context of observation. The basic problem is how perceptual strategies have developed (Payne, 1982). Further, if perceptual strategies are as a result of pre-conscious automatic processes they should be invariant across subjects - given common experience. Prospect theory does accommodate findings of risky choice studies but not others (Payne 1982). The theoretical models described so far have attempted to extend the role of normative, prescriptive models to that of a descriptive model. Other work has been concerned with the way in which algebraic models can adequately describe the differential use of cues by decision makers.

Shanteau (1972; 1975) proposed an algebraic model based on Anderson's integration theory. Anderson (1972) used a regression approach to develop a model that described the way people weight and combine information. By using a simple differential weighting of cues, Anderson described configurality present in clinical judgement that had previously been observed by standard regression techniques. Goldberg (1968) had concluded that a linear, additive model appeared to predict the judgements made by doctors and clinical psychologists quite well. Anderson pointed out that a direct application of regression techniques to the study of judgement is misleading because they obscure configurality that may be present in the data. A simple regression technique depends on additive, equally weighted beta parameters and can impose linearity on data that may not be linear. It does not allow the person to differentially weight different signs or symptoms, seeing some as more serious or significant than others. Anderson proposed a differential weighting and averaging model which successfully described configurality in judgement which had been obscured in the linear model. This can be seen as a good surface description of the decision/judgement process but the underlying cognitive processes may be rather different to the surface form of the model as has been indicated by other work reviewed. Wallstan (1972; 1977) has formulated a model similar to that of Anderson (1972) and assumed that people process cues for the purpose of evaluating one hypothesis relative to another. Which in a natural decision situation may be an over simplification, where more than two possibilities may be under consideration by the decision-maker.

A very different model was proposed by Abelson (1976). He suggested a model of script processing based on the concept of a "cognitivescript", a cognitive script being "a coherent sequence of events expected by the individual involving him either as a participant or an observer". A script in this case would be a diagnosis stored with the associated signs and symptoms, and the expectations of occurrence based on previous experience and learning. It is possible to postulate a role for both the "representativeness" and the "availability" heuristics, where a script may be accessed. Either because the situation in which the decision is being made is sufficiently similar to a previously encountered situation, or because of a judgement based on frequency.

The process by which Abelson suggested cognitive scripts are formed will be described, followed by the theory of how people use them to make decisions. Firstly, a vignette recording a single episode may be stored as a single unit and this can happen repeatedly. It is suggested that similarity groupings can then build up categorical vignettes from many single experiences of a given type of situation. The similarity groupings are said to occur as a result of a feature abstraction from a single vignette and a pattern recognition decision being made on the basis of feature matching or similarity. It is still unclear as to how people do assign objects or experience to categories, whether it is by a matching of features (Selfridge, 1959) or by a more abstract description of the structural functional dimension of an object or experience (Sutherland, 1973). Abelson suggested that with enough experience in a single domain, lists of features can be processed instead of each vignette that makes up the script, ie. the script as a whole is processed.

Abelson suggested that in natural decision-making, it is clear that statistical decision models are inadequate to reflect actual cognitive processes or predict outcomes. People do not ordinarily make decisions based on combinations of prognostic cues (Dawes, 1976). In terms of decision-making, Abelson suggested the three levels at which scripts are formed reflect three levels of later use. First is the episodic level where a past, single case recalled is similar to the present case. Then there is what is called the categorical level, where a generic type is involved, eg. stereotype of a particular patient group. Lastly, there is the abstract/hypothetical level which consists of bundles of pro's and con's, where success is a hypothetical variable contingent on all important innumerable features. It is at this level that the work described until now has assumed to be the level at which people operate when making decisions.

Abelson suggested that this may not be the case. He conjectured that they are predisposed to operate at the first and second categorical levels in many instances. The "availability" bias would suggest the first level and the "representativeness" bias the second.

The model can be seen to be concerned with information acquisition, the way it is processed and the resulting output. Abelson (1981) described a script as nothing more or less than the simultaneous activation of a set of events, any of which may become involved in local inference. But there are differing levels of inference and not every simple inference involves a script. Therefore, for a script to have status as a cognitive structure, it must embody more than a simple inference rule.

There is good evidence of a simultaneous activation process (Anderson, 1980). The priming of one script event from a story lead to especially fast recognition of another script event from that story. The priming of the script name led to faster recognition of script events (den Uyl & van Oostendorp, 1980). A major test of the power of scripts to organise understanding was the gap-filling phenomenon found in studies of prose comprehension. People have shown a strong tendency to falsely recognise non-mentioned script events (Bower, Black & Turner, 1979; Graesser, Woll, Kowalski & Smith, 1980). This effect was discussed earlier in relation to the work of Arkes & Harkness (1980) who found that the effect of making a diagnosis was to falsely recognise related symptoms. Further, the strong preference shown to maintain correlation between symptoms suggested a script-like organisation (Medin, Altom, Edelson & Freko, 1982).

Convergent work concerned with processes of social categorisation can lend indirect support for the notion of a script. Stereotypes (Tajfel, 1978), "schema" (Markus, 1977: Taylor & Crocker, 1981), implicit personality Theories (Schneider, 1973), "prototypes" (Cantor & Mischel, 1977) and scripts can all be seen to be conceptually very similar. That is when the script is being used at the first or second categorical level. Social categories can lead to encoding biases (Hamilton, 1976: 1979: Hamilton & Rose, 1980). That is they can determine the acquisition stage or processing and subsequently encoding by determining what information in the environment or about a person is selected ie. setting up prior expectations that are then sought to confirm. Behavioural consequences resulting from bias have also been observed (Snyder, Tanke & Berscheid, 1977).

One of the major problems for a theory of decision making is the lack of performance invariance characteristics of much of the work. This has made the identification of a small number of underlying principles that can describe decision behaviour complicated (Payne, 1982). Abelson's model allows different strategies to be incorporated at different levels. The problem arises as to how the level of decision performance is determined and to define precise mechanisms and processes that are encompassed within an "all embracing" theory that can predict strategy use.

1.7

Studies of Nursing/Medical Decisions

Clinical inferences made by nurses first became the subject of study in the mid-sixties with a series of studies that looked at, amongst other things, the information units used by nurses and the search strategies that they employed (Hammond, Kelly, Schneider & Vancini, 1966; Hammond, Kelly, Castallan, Schneider & Vancini, 1966). The results were disappointing and this was largely due to a lack of understanding of natural decision making and an assumption of statistical reasoning.

Medical decision making has received considerable attention (Elstein & Bordage, 1979). A major focus of decision improvement has been computer-based aids (Shortlife, Buchanan & Feigenbaum, 1979). It is intended to give a very brief overview of work concerned with natural medical decisions and then consider a study of nurses' decisions. Many studies of medical decisions have been directed towards medical education. Expert performance has been analysed identifying aspects of that performance to be emphasised when teaching students (Elstein, Shulman & Sprafka, 1978). Other work has been concerned with attempting to identify the organisation of factual knowledge and inference rules necessary for effective clinical decisions (Kassirer & Gorry, 1978). As has already been discussed, the consequencies of making a diagnosis for later symptom pairing and recall have been studied (Medin, Altom, Edelson & Freko, 1982; Arkes & Harkness, 1980); and the interpretation of medical information has been found to vary as a function of experience (Johnson, Hassebrock Duran & Moller, 1982). Yet, Elstein, Holmes, Ravitch, Rovner, Holzman & Rothert (1983) in a brief review of work in this area, noted that it still remains to integrate the results of studies of medical decisions into a more general theory of information processing.

Elstein et. al. (1978) adopted a problem-solving approach to examine clinical reasoning. They sought to describe the thoughts and steps that characterise a doctor solving a clinical problem in a paradigm that is more representative of the medical setting. The method employs verbal protocols and direct observation. The transcripts are analysed and from this an account of reasoning is constructed. This paradigm exercises less control over the type of information sought and collected than is true of a normative approach. Generalisation is not from a tightly controlled (but artificial) research setting, rather account is taken of clinical reality where problems are more "open" and less well defined.

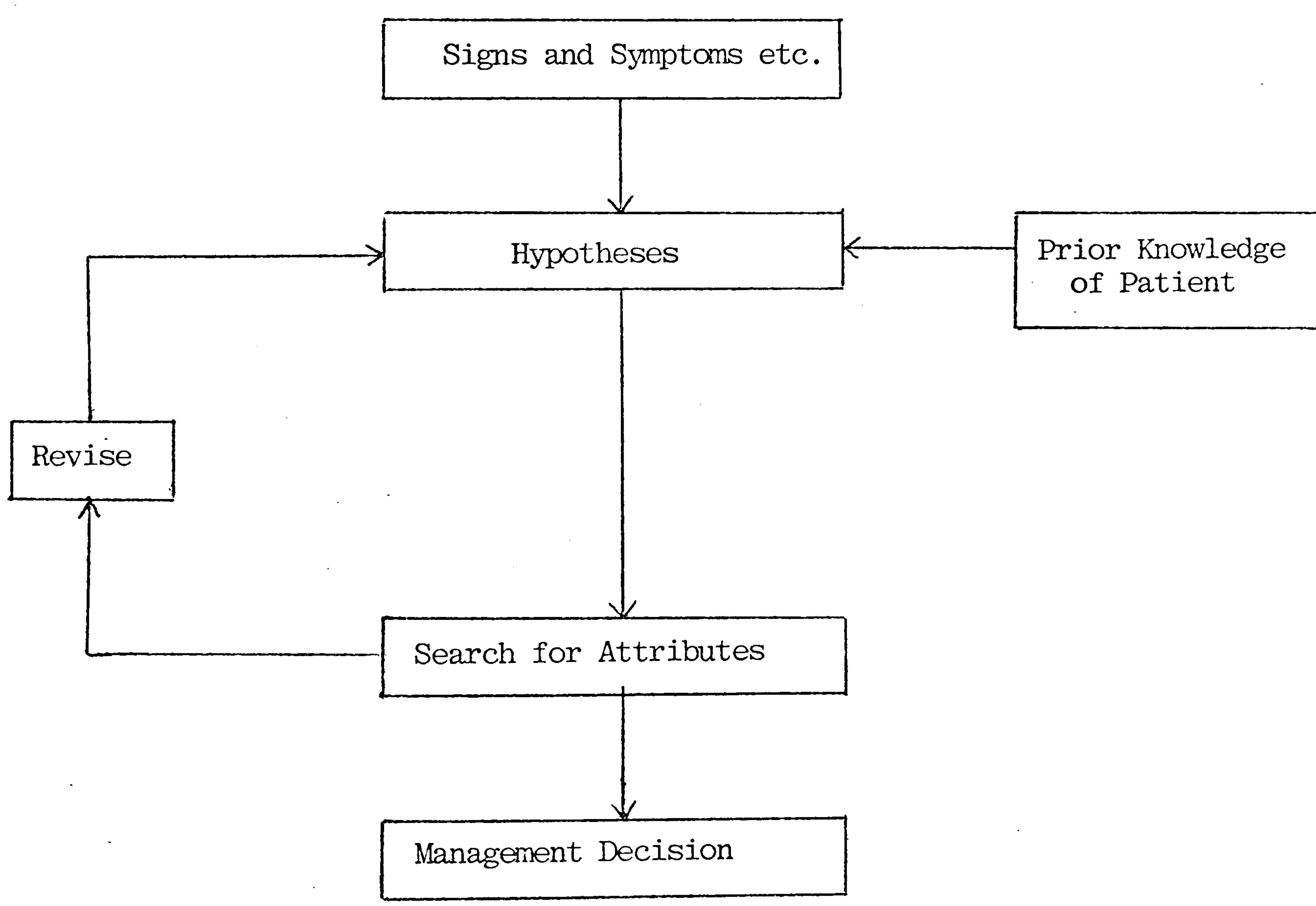


Fig.3 Stages in diagnostic problem-solving (from Weinman, 1981)

Elstein et. al. (1978) found that clinicians engaged in diagnosis commonly employed the strategy of generating and testing hypothetical solutions to the problem. They formulated hypotheses early in the task and these were used to guide the search for further information and an acceptable solution. This was contrary to accounts of diagnostic reasoning that assumed that the clinician amassed a large body of information prior to forming any diagnostic hypotheses, which were suggested by the overall pattern of that information. The stages in the problem-solving process that were found to characterise medical problem-solving are illustrated in Figure 3. The initial presenting symptoms or complaints led to the generation of hypotheses . The choice and ranking of these hypotheses was influenced by several factors. There was the clinician's orientation (eg. psychosocial), considerations of disease incidence, the seriousness and treatability of the disease, and personal knowledge of the patient and the patient's background. All these factors were found to contribute to the initial formulation of hypotheses. This stage was followed by a search for attributes such as signs, symptoms and test results to support or refute these hypotheses. But clinicians have been found to have difficulty testing negative attributes (McWhinney, 1972).

Given the natural limits imposed on processing capacity, this reasoning process transforms the ill-defined "open-ended" problem of "What is wrong with the patient?" into a series of better defined problems (Elstein & Bordage, 1979). By constructing a set of hypothesised end-points (diagnoses), it becomes possible for the clinician to work backwards from the diagnostic criteria suggested by each hypothesis to the examination to be conducted. The search for information is simplified because only certain items will be sought. Elstein et. al. (1978) observed that hypotheses were retrieved using very few cues and the number of hypotheses considered simultaneously was usually limited to four or five. It was the salience of cues that was considered critical to retrieval because of the suggestion that "availability" as subjective probability was experienced as vivid, salient possibities (Tversky & Kahneman, 1973).

The theoretical perspective adopted by these studies was the information processing or problem-solving model of Newell & Simon (1972). The tasks employed by Newell and Simon were moderately difficult problems of an abstract nature, including chess, symbolic logic, and algebra-like puzzles. From this work an essentially dynamic model was developed, which assumes a limited capacity for rational thought. The model suggests that hypotheses are generated and tested in order to work to a solution. Central, is the theoretical reduction in size of the problem space, and changes in the representation of a problem can change the solution process significantly. The first step in solution was suggested to be the development of an understanding of the nature of the problem and deriving an internal representation through the processing of key features. At this stage an understanding of the nature of the goal, particular conditions imposed and the type of information available is established, and as a result plans for solution are made. The distinction is made between alogrithms and heuristics. Newell & Simon suggested a predominant reliance on heuristics or strategies to find direction and guidance to solution in many situations. Critical is the generation of hypotheses of how a solution can be obtained, and the device and use of strategies to that end. The conversion of an "open" problem to a "closed" problem where the goal is defined as one of a limited set of possible goals (eg. diagnoses) reduces the number of attributes or items of information requiring testing. This would predict a limited search, and revision of hypotheses would predict a longer search.

The problem-solving process was suggested to progress from one state of knowledge to another by the application of a small set of heuristics. If new information is obtained, progress is made to solution. But, where no new information is obtained and an impass is reached, revision takes place. A common and powerful heuristic is to find analogies between the present problem and those experienced in the past.

Gordon (1980) studied nursing diagnosis and took a different approach to that of workers engaged in decision analysis adopting the approach of early decision theorists (Bruner, Austin & Goodnow, 1956). Gordon suggested that, given processing limitations a "risky" but manageable strategy in a diagnostic task was to reduce the total number of possible diagnostic hypotheses to a manageable set, the alternative being to test every possible diagnostic hypothesis sequentially until a decision was reached. The first strategy implied an initial collection of information that was high in information content and allowed a discrimination at the start of a task as to the most likely diagnoses. The distinction was made between general, relatively static information that could serve to place the patient in context eg. age, type of surgery and more specific, relatively changeable information such as signs and symptoms.

The nurses studied were a group of sixty Master of Science in nursing graduates and their task was the diagnosis of a post-operative complication. Two complications were used, atelectasis and haemorrhagic shock, from a possible six. A list of these six complications, in all possible combinations, making a total of 32 possible patient conditions, were given to each nurse. The nurses were instructed that their task was to determine the current condition of a patient. The only information given was that the patient was post-operative following general sugery. A selection paradigm, using a modification of the game of Twenty Questions, was employed. Subjects were able to control the type and sequence of information that they obtained by asking for each item singly. Each request for information was treated as a trial and provided information about the attribute being tested and the number of hypotheses the nurse said that she was testing. The statement of the diagnosis reached was made at the end. The two patient conditions were a condition of unrestricted information, where atelectasis was always the post-operative complication, and a condition where item requests were restricted to twelve and haemorrhagic shock was always the complication.

The main findings were, in the first half of the task, significantly more multiple hypothesis testing took place, more so in the restricted condition ($p<0.001$) than in the unrestricted condition ($p<0.01$). People used more general items to set the scene and provide a context. Further, inaccurate subjects in the unlimited condition used more predictive testing in the second half of the task ($p<0.001$) than subjects who were accurate. These findings imply that predictive testing early in the task set the scene, ie. the use of general information to reduce the number of hypotheses provides an effective strategy early in a diagnostic task, especially when a person is restricted due to constraints of time.

There was, however, a difference in performance between the two diagnostic tasks performed. In the unlimited condition, where people could request as many items as they felt they needed, 48% of the subjects attained a correct diagnosis. But where people were limited to twelve items, 88% of people were accurate. This was a surprising finding, from the preceding review, it would have been reasonable to predict that the accuracy of performance on the two tasks would be equivalent. Higher confidence levels would have been expected for the unlimited condition because of the association of confidence increasing with the amount of information (Tversky & Kahneman, 1973). However, the mean confidence level in the unlimited condition was 77% and in the limited condition it was 81%.

The differential levels of accuracy could have several explanations. The design of the tasks could have imposed excessive strain on memory in the unlimited condition.

People were required to recall the items that they needed to know from memory. Then they needed to evaluate and retain the information that they were given. It is possible that where the information was restricted, people organised information more efficiently. However, a problem of experimental control does make any interpretation equivocal. In the unlimited condition the correct diagnosis was always atelectasis, and in the limited condition it was always haemorrhagic shock. There is no reason to assume that these diagnoses were equivalent in difficulty, therefore task difficulty did confound the observations.

1.8

Aids to Decision Making

Aids to medical decision making have largely concentrated on computer based techniques. The potential of decision analysis and analytic epidemiology has been recognised for some time (Beck, 1983). This includes both clinical medicine and medical education, but acceptance has been slow, although software (programmes) has been becoming available allowing a variety of different analyses. Beck (1983) noted two hindering factors, the first being a lack of accessible computer facilities and secondly a lack of statistical skills among physicians. The lack of statistical skills is reflected by an underuse of Bayes' theorem which only requires a pocket calculator.

Bayes' theorem has been used and has been shown to perform as well, if not better, than the clinician (Slovic et al. 1977). Einhorn (1972;1974) demonstrated the effectiveness of "bootstrapping" expertly defined and weighted variables with an algebraic model of that weighted policy. This showed how expert judgement and statistical techniques can incorporate poorly defined and hard to measure variables into a judge's model. With the advent of the microcomputer a variety of decision analyses can be more conveniently employed. Decision tree programmes that allow probabilities and utilities to be calculated and sensitivity analyses (Corrections for uncertainty) are available and relatively simple to master (Silverstein 1983; Lau, Kassiter & Parker, 1983). These methods allow risk analyses of alternative causes of action to be calculated, they are not simply confined to aiding diagnosis.

Another development has been that of expert systems where the machine is programmed with sufficient information about a clinical area, plus a mode of decision analysis to simulate the role of an expert in that clinical area. The first programme of this type was MYCIN which was developed for use by clinicians who were not experts in the field of bacteriology or pharmacology to aid their correct use of antimicrobial therapy (Shortliffe, 1976). This work is in the early stages of development but holds promise for the future in terms of making expertise available to non-expert practitioners.

The techniques that have been employed to aid medical diagnosis and decision making can be seen to be the result of evidence that natural decision makers can positively benefit from employing statistical models when combining information from different sources. But, useful as this approach may be for many medical applications, it can be seen as yet to have limited application for nursing. There are several problems. Firstly the use of relatively sophisticated quantitive techniques has appeared to daunt many medical practitioners (Beck, 1983). Nurse training does not develop necessary skills nor do many nursing decisions warrant that degree of sophisticated intervention. But, the development of the use of information technology in intensive care units offers the possibility of investigating machine decision aids for nursing decisions in the future.

Decision aids employed in management and business mostly rely on taking a problem, such as identifying a potential market, and decompose the total problem into a series of structually related parts. The decision maker provides subjective assessments of the component parts (Slovic et al. 1977). Humphreys & McFadden (1980) suggested that the usefulness of decision aids may be as a result of structuring the task and consequently clarifying the nature of the goals ie. helping people to evaluate information that has already been acquired. Russo (1977) found that when unit prices were presented to supermarket shoppers in organised lists that were ordered by the relative size of the unit prices, thus giving a simultaneous presentation of information, purchasing patterns were changed. The patterns were changed for people who had this information compared with shoppers with no information or when the information was simply indicated next to the products on the shelves (implying sequential acquisition). This decision aid, instead of helping people to evaluate information that had already been acquired eg. through "bootstrapping", reduced the strain on memory and attention by aiding the acquisition process. But a better understanding of attention and memory processes is necessary for the approach to predict success (Einhorn & Hogarth, 1981).

Decision-tree formats have been widely applied. Fischhoff, Slovic & Lichtenstein, (1978) presented expert and novice judges with diagnostic check-lists represented in a decision tree form known as fault trees. The diagnosis was the cause of mechanical failure of a car and the results indicated that the apparently comprehensive format of the fault tree blinded both groups of judges to the possibility of missing causes of malfunction. Einhorn & Hogarth (1981) emphasise the role of attention and memory in the decision process and also discuss the relevance of context. They noted that context had typically been defined in terms of task variables and they emphasise the importance of the decision maker's past experience and learning, plus the cognitive apparatus. The role of context that is set by past experience has been curiously overlooked or ignored in much of the work in decision making.

The following section presents a summary of the influences on the natural decision process in the stages of the general model presented in Figure 1.

1.9

Summary of Some of the Biases Operating at Different Stages of the Judgement Process

	BIAS	DESCRIPTION
Acquisition of Information	AVAILABILITY	The ease with which instances can be remembered affects judgements of frequency
	SELECTIVE PERCEPTION	People seek and anticipate information that is consistent with their own hypotheses, often ignoring disconfirmatory evidence
	FREQUENCY	People judge the relationship between a cue and criterion by its observed frequency rather than by its observed relative frequency
	"CONCRETE" THINKING	People tend to ignore base-rates or information about prior probabilities
	ILLUSORY CORRELATION	The belief that two variables co-vary when they do not
	DATA PRESENTATION	The context in which an event is observed can affect the perception of that event. Order effects of primacy/recency where the first or last cues encountered assume undue importance
Processing of information	CONSERVATISM	Decisions that are more regressive than Baye's theorem would predict
	REGRESSION BIAS	Extreme judgement made with insufficient evidence
	ANCHORING & ADJUSTMENT	Decision made by anchoring on a cue and then adjusting for new information
	REPRESENTATIVENESS	Likelihood that two people belong to the same category judged by the degree of similarity of essential features

	BIAS	DESCRIPTION
PROCESSING (cont.)	"BEST-GUESS" STRATEGY	When there is much uncertainty, people base decision on most likely hypothesis
	GORDON'S CONTEXT ITEMS	Less accurate nurses used predictive context items late in the decision. Could be "best-guess" or lead to cognitive strain
	COGNITIVE STRAIN	People unable to do simultaneous processing required to gauge probabilities accurately
	<u>DECISION ENVIRONMENT</u>	
	COMPLEXITY	People unable to experimentally control and vary elements in the environment - cognitive strain
	STRESS	Emotional stress can lead to panic judgements (Janis & Mann, 1977)
	CONTEXT	Task variables plus the decision maker's past experience and learning
OUTPUT	SCALE EFFECTS	Scale on which responses recorded can affect responses (Slovic & Lichtenstein, 1971)
	ILLUSION OF CONTROL	Activity concerning an uncertain event can in itself induce feelings of control over that event (Langer, 1975)
FEEDBACK	ACTION	Effect of taking an action destroys the likelihood of the equal occurrence of the alternative outcome
	OBSERVED DECISION	People observe positive hits more than mistakes
	MAKING A DIAGNOSIS	Has the effect of later falsely recognising cues that may not have been present at the time

CHAPTER 2

2.1 Rationale for The Experimental Work

The objective of the first part of the work reported within this thesis was to attempt to identify a means of helping nurses to make more accurate decisions. Several points emerged from the work discussed in the review. It was found that feedback or information about the outcome of a decision had little or no effect on the accuracy of subsequent decisions. Nor can people be turned into normative decision makers. Computer applications were discussed in relation to their use in medical decision making, but the usefulness to nursing decisions is, as yet, uncertain for several reasons. First there is a lack of knowledge available about the information that nurses use to make many of their decisions and secondly, computer techniques may not be appropriate for many, apparently less complex decisions.

Decision-tree formats, although widely applied, have been found to obscure possible cues or variables because of their apparent completeness (Fischhoff, Slovic & Lichtenstein, 1978). Einhorn & Hogarth (1981) noted the lack of attention paid to the role of context set by past experience and learning in much of the work in decision making. Although, by contrast, an interesting observation was that of Russo (1977) who found that aiding the order of acquisition of related information helped people to make more use of that information.

A diagnosis has been proposed to be analogous to a schema stored in memory along with associated cues. Among the cues stored with any diagnosis are those that relate to the context in which a particular diagnosis may be expected. Therefore, a diagnosis can be viewed as a coalescence of signs and symptoms, patient characteristics and the context in which they could occur as well as the category/diagnostic label. All the evidence reviewed indicated the difficulty in isolating a small number of principles underlying decision making. But one general observation has emerged that suggested that the strategy people employed when making a decision was one that made the cognitive task manageable ie. tended to simplify the task.

An examination of the totality of the processes involved in making a decision has been shown to be a very complex exercise. It was the initial phase of the process and the consequences for the subsequent phase that was examined in this work. This could be of more interest to nurses in terms of patient problem identification and a specific method of planning care eg. the Nursing Process. It was of interest to investigate how nurses set about searching for information when making a clinical decision ie. identifying a problem.

Bruner, Austen & Goodnow (1959) identified two potential search strategies that are available when making decisions. In the first instance the decision maker could hold in memory all potential clinical judgements, and when all the available information has been collected select a judgement that fits most closely. Alternatively the decision maker could start out with a small set of likely judgements based on initial information, then search for evidence to test them. The first strategy is very difficult as it imposes considerable strain on the cognitive apparatus. The second requires an initial selection of information that allows an identification of likely judgements, and consequently the amount of information that needs to be manipulated. When deciding about the current condition of a patient, information that serves to set the patient in a context can help to reduce the number of judgements needing to be examined. This would be general, background information about that patient. Kassirer & Gorry (1978) showed that specific diagnostic hypotheses were generated often with little more information than the patient's age, sex and presenting complaints.

Gordon (1980) adopted a similar approach although there were problems of task difficulty. One diagnostic task induced a ceiling effect. But, nurses did initially select this type of information more often at the beginning of a diagnostic task than later. Further, when the nurses could ask for as much information as they required, those who asked for the general information later in the task were less accurate. This suggested that it was the initial use of general, context information to predict the relevance of one hypothesis versus another at the beginning of a task that served to make the task orderly and manageable, but predicting from general information at a later point served to confuse. Almost as if another diagnostic schema was accessed with its associated information and caused confusion. These nurses were very highly educated in surgical nursing and had considerable experience, they would have been familiar with the patient condition that required diagnosis. The more accurate nurses demonstrated an orderly strategy that made their task manageable, it was considered to be of interest to investigate whether less experienced nurses did the same. In the U.K. the nurses with the closest contact with patients are inexperienced learners supervised, in many situations, by recently qualified staff.

It was found by Russo (1977) that aiding information acquisition can facilitate the use of that information. But a better understanding of memory and attention processes is necessary to predict success (Einhorn & Hogarth, 1981). The paradigm that Gordon employed required the nurses to rely entirely on their memories for both recalling what they wanted to know and remembering the answers given. It is unclear whether nurses were inaccurate because of an inefficient search strategy or because of an excessive retentive load on memory. It is also unclear as to whether similar findings would be observed when the retentive load is removed. Given that search/attention and memory, although inter-related, make differing contributions to the total decision process, attention serves to facilitate recognition of the relevance of information that is being subject to manipulation where there appears to be a limited processing capacity (Baddeley & Hitch, 1974: Kahneman 1973). Further, in the natural decision environment much information is provided, such as general patient characteristics, signs and symptoms. What is required is frequently the recognition of the relevance of this information.

The possibility that was considered to follow from this was that if inaccuracy was further found to be associated with a use of general information later in the task rather than predominantly at the start when the effects of memory and attention were disassociated, this could be potentially useful. Nurses could be taught to use context information at the start of a decision task. But, the use of general, context information later could serve to predict fresh possibilities where initial hypotheses had not been confirmed, where a retentive burden is removed.

The evidence reviewed did indicate the use of a "representativeness" heuristic in many judgement situations. This heuristic is closely connected to the strategy discussed and it was considered to be of interest if some people were more predisposed to make use of this heuristic than others, and whether this could be implied by their level of cognitive complexity (Bieri, Atkins, Briar, Lobeck, Miller & Tripoldi, 1970). Cognitive complexity refers to the number of dimensions that a person uses to describe their social world. People are considered to be relatively complex if they use many dimensions. It was speculated that a person who was more complex would use more general information to evoke a more complex background context.

It was decided to partially replicate and extend Gordon's experiment using a recognition paradigm and incorporating an intervention designed to induce the early use of context information. Levels of cognitive complexity were also measured. The next chapter describes pilot work in which the materials were developed and the intervention was tested for feasibility.

2.2 Pilot Study

The procedure used by Gordon (1980) will be briefly outlined. Subjects were Masters graduates in surgical nursing. They were instructed that their task was to evaluate the condition of a patient who had recently undergone general surgery. The diagnoses were post-operative complications. They were called upon to recall from memory the information that they needed to know. They asked questions to request information serially and the type of information was recorded ie. general background and relatively stable information and specific sign and symptom information. The hypothesis that the subject reported to be testing was also recorded. In one task they could ask as many questions as they wished, in the other task they were restricted to twelve questions.

This pilot study employed a recognition paradigm and did not restrict the amount of information that could be requested, included the intervention and a measure of cognitive complexity. The hypotheses were not formally proposed at this stage. The concern was the feasibility of the materials and the intervention.

Method

Design

The independent variables were:-

1. Type of information available
    a) General information
    b) Specific information
2. Experimental intervention
3. Cognitive Complexity

The dependent variables were:-

1. Accuracy
2. The number of questions of the different types of information

General information was defined as relatively stable facts about the condition of a patient.

Specific information was defined as changeable facts about the condition of a patient.

The task was to diagnose the current condition of a general surgical patient

The only information given was that the patient had received general surgery. The general surgical condition selected was appendectomy, a commonly occurring surgical procedure. The conditions for clinical diagnosis were three post-operative complications:-

A - Paralytic Ileus

B - Pulmonary Embolism

C - Urinary Retention

The sequence of presentation was partially balanced as follows:-

Patient Condition	Subjects
AC/B	6
CA/B	6
AB/C	5
BA/C	5

where two tasks, the first being for practice were followe by the intervention and the third task as control for task difficulty and practice.

Questions were asked serially and were asked on the basis of recognition of an appropriate item from an item list.

The intervention was an illustrated lecture, pre-written to standardise presentations between the groups (App 1).

SUBJECTS

Twenty-two student nurses at the end of their second year of training at a London Teaching Hospital. The age range was 20-32 years. The population was female with a range of educational achievement from five G.C.E 'O' levels to nine 'O' levels and three 'A' levels. This comprised one set of student nurses who had completed their surgical module and had all had surgical nursing experience. The surgical module contained formal teaching about post-operative complications and the associated pre-disposing states, plus signs and symptoms.

MATERIALS

1. A two-part questionnaire (App 2) assessed:-
   a) Personal details, ie. age, sex and educational qualifications
   b) Cognitive complexity measure (Bieri et al. 1970)
   Four of the roles were altered to reflect a hospital setting ie:-

   ROLE

   4. A patient you'd like to help
   8. A patient with whom you feel most uncomfortable
   9. Ward Sister
   10. Patient difficult to understand

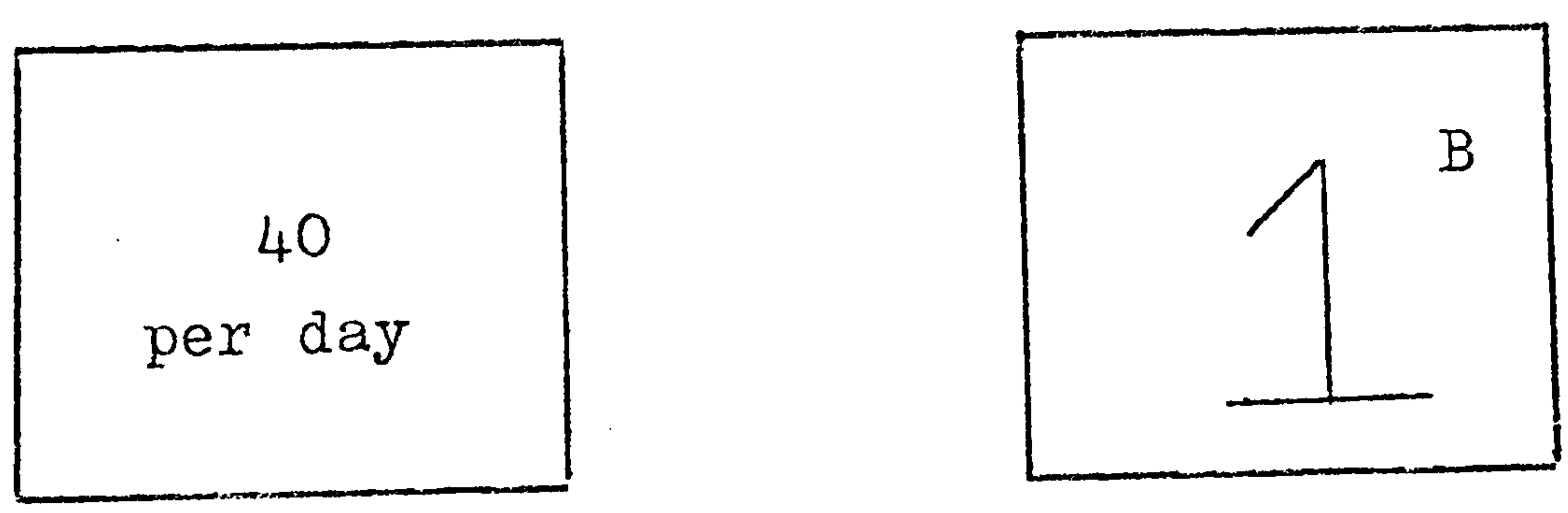
2. Standardised lists of questions for requesting information (App 3). An example of the layout being:-

Question:- What or how much?			
Card	Question	Order	Answer
1.	Cigarettes smoked daily		
2.	Pain now		

The general and specific questions were blocked in logical groups, eg. all vital sign information appeared sequentially and the blocks were randomised. There were a total of 39 questions. The information about the sign and symptom clusters associated with the chosen complications was taken from a standard text known to be used as ready reference by this population of student nurses (Fream, 1978 "Notes on surgical nursing, App. 4). Only one diagnosis was appropriate for each task.

3. A list of nine possible patient conditions, including the three target conditions, in randomised order (App. 5).

4. Packs of 39 individual "playing" cards with the answer to a request on one side and the sequence on the reverse, eg. for Question 1 above, condition B.



Three packs, one for each diagnostic condition

5. A pre-written lecture with transparencies instructing the subjects to use general background information initially to set the scene, and then look for signs and symptoms.

Procedure

The session was conducted in a classroom of the School of Nursing of a London Teaching Hospital. Questionnaires bearing the subject number had been placed on each desk, next to the appropriate subject number, face down, prior to the arrival of the subjects. When they arrived, the subjects were asked not to look at the questionnaire immediately. The researcher then introduced herself and gave a brief explanation of the research project in which they were about to take part (App. 6 General Instructions). The intention of investigating reasoning strategies that could be taught was explained.

Reference was then made to the questionnaire and the subjects were told that this would be followed by two short tasks, instructions for the task being given after questionnaire completion. Instructions were given about not writing names on any of the test materials. No mention was made at this point of the intervention or the third task. The subjects were then asked to complete the questionnaire, not pausing for too long on any one item, having first read the instructions. The questionnaires were collected on completion.

Two packs of cards were then given to each subject, plus two lists of information requests and a list of possible patient conditions. The packs and lists had previously been ordered and marked for the appropriate running order for individual subjects. Instructions were then given as to how to proceed with the tasks (App. 7). The subjects were asked to imagine that they were caring for a general post-operative patient. Their task was to assess the current condition of that patient. That condition would be one of the conditions that appeared on the list they had been given, and for each task the condition would be different. It was explained that in order to do the task, they should take the first list and pack. Then select from the list the item of information that they required. They were told to take from the pack the card bearing the appropriate number. Then to look on the reverse side and the information they required would be available. They were asked to record the order in which the card was selected, and the information contained on it in the appropriate columns on their lists. Repeating this procedure with each item until a diagnosis was made.

The subjects were asked not to look at the task being done by their neighbours or consult with each other. they were told that they could start anywhere on the list, there was no logic in the order and asked not to start with Item 1 first, unless they really wanted that piece of information. It was stated that not all the questions would be relevant to the task that they were doing and that they would probably need to ask about one-third of those available. When they had finished, the diagnosis reached was recorded in the space provided. Each subject did the first task, followed immediately by the second. When they had finished all the test materials were collected.

The experimental intervention was then given by the researcher, accompanied by overhead transparencies. The intervention was in the form of a pre-written lecture (App. 1) which attempted to provide a strategy for making a diagnosis. The types of information were related to a script theory of human memory where diagnoses are postulated to be Gestalts of all the relevant information items. This was followed immediately by the 3rd task. Subjects were then thanked for their co-operation and told the identity of the task diagnoses. They were asked not to discuss what they had done with any of their colleagues as they could be possible subjects of further sessions.

Results

Table 2.1 Table Of Means Of The Number Of General And Specific Items Used For All Subjects

TASK	1st ½ Task General	Specific	2nd ½ Task General	Specific	
1	2.5	4.1	1.1	6	13.8
2	2.04	3.5	1.04	5.04	11.6
3	3.2	3.2	0.8	5.9	13.1

N = 22

As can be seen from Table 1, more general questions were asked in the first half of the task compared with the second half.

Of the total 22 subjects, only one subject was inaccurate in the first task, three were inaccurate in the second task, and two in the third task. Therefore, no analysis were performed. The "ceiling" effect present in this data, and possible reasons for it, will be discussed in the following section.

A X^2 analysis compared subjects who asked a total of 12 questions in Task 2, with subjects who asked < 12 questions. The pieces of general information considered were type of surgery and length of time since surgery. It appeared that subjects asking ≥ 12 questions were significantly more inclined to ask these general questions at the beginning of the task ($X^2 = 5.32$: $p < 0.05$ App. 8). A X^2 on the same basis failed to show any significant differences for Task 3 (App. 9).

All subjects successfully completed eight of the ten roles of the cognitive complexity questionnaire. However, a proportion of the subjects found difficulty with the roles:-

8. A patient with whom you feel most uncomfortable
10. A patient difficult to understand

No analyses could have been performed with this measure in view of the findings that indicate that when making negative decisions people use more information (Wallstan, 1977; Morlock, 1967). The negatively toned roles imply that more constructs could be used by most people, therefore any overall mean score would be distorted.

Discussion

This pilot study demonstrated that the diagnostic tasks, as they stood were not only within the subjects' range of competence, but were too easy. This ceiling effect that was present in the data indicated a need for modifications to make the task more difficult. The running of the session was the first area in which alterations needed to be made. Subjects did whisper throughout, therefore, the range of possible diagnoses could emerge, and as the blocking procedure did not control that, people sitting next to each other should have done different tasks, it was possible to copy. For future sessions it was intended to adjust the blocking so that no two neighbours ever did the same task, and to run the session in silence. The experimental instructions could be modified to imply a wider range of possible diagnostic tasks being attempted by subjects. The list of possible diagnoses could be extended, therefore, making a guessing strategy less effective.

As can be seen from Table 1 in the results section, a common strategy in the pre-intervention task was to use general items more in the first half of the task than in the second half. An X^2 analysis of task 2 revealed that people who used fewer items, ie. ≥ 12, used different pieces of context information. The subjects used the items "type of surgery" and "length of time since surgery" significantly more often ($p < 0.05$) in positions 1, 2 and 3 than did subjects using < 12 items. This was not apparent in the third post-intervention task.

There were problems with the measure of Cognitive Complexity for reasons mentioned earlier. The negatively toned roles, patient with whom you feel most uncomfortable and a patient difficult to understand, was not completed by a proportion of subjects. The finding that when making negative decisions people use more information (Wallstan, 1977; Morlock, 1967) implies more complex use of constructs as a general effect. This would make any compensatory method taken for the loss of two scores, eg. taking a mean for all roles completed and multiplying by the total number of roles, biased toward relative cognitive simplicity.

The modifications intended to tighten the design for the next pilot were:-

1. Alternation of balanced conditions so that people sitting beside each other did not do the same task

2. Run the session with instructions that asked people not to talk at all

3. Extension of the list of possible complications from which to choose

4. Revert to the original format of Bieri et al. (1970) in the cognitive complexity questionnaire

2.3 Pilot Study 2

Method

Design

As for Pilot 1, except for the balancing of the sequence of presentation.

Partial Balancing Order

Patient Condition	Subjects
AB/C	8
AC/B	7
BA/C	7
CA/B	7

Where subjects were allocated in serial progression to the four different orderings.

Subjects

29 Student nurses at the end of their second year of training at a London Teaching Hospital. The age range was 19-28 years, the mean being 20.8 years. 27 subjects were female, 2 were male. 5 subjects were graduates, the remainder were educated to 'O' or G.C.E. 'A' Level.

Materials

Were unchanged, apart from the extended list of possible complications:-

1. Peritonitis
2. No complication
3. Pulmonary embolism
4. Haemorrhagic shock
5. Chest infection
6. Paralytic ileus
7. Urinary retention
8. Thrombophlebitis
9. Deep vein thrombosis

10. Septicaemia

11. Atelectasis

12. Wound infection

13. Reduced blood pressure

14. Asphyxia

15. Burst wound

16. Hernia

17. Pelvic abscess

18. Subphrenic abscess

19. Iliac abscess

20. Adhesions

And the cognitive complexity roles which were now:-

8. Person with whom you feel most uncomfortable

10. Person difficult to understand

4. A person you would like to help

PROCEDURE

The procedure was identical to the first pilot study, with the exception of the inclusion of specific instructions for silence and no collaboration. "Please do these tasks in complete silence, I do not want you to discuss anything with your neighbour, even when you have finished.....People sitting alongside you will never be doing the same task...."

RESULTS

The subjects were not distinguished between on the basis of level of educational attainment. One of the five graduates performed all three tasks correctly. One got 2 out of 3 correct, and 3 got 1 out of 3 correct. This is similar to the group's performance as a whole.

TABLE 2.2

		TOTAL MEAN	S.D	CORR. MEAN	INCORR. MEAN	MEAN ITEMS
TASK I	1st ½ % general	35.4		37.0	33.3	
	% specific	64.6	17.0	63.0	66.7	
	2nd ½ % general	14.9		14.9	16.4	12
	% specific	85.1	12.6	85.1	83.6	
				N=17	N=12	
TASK II	1st ½ % general	45.0		46.6	42.0	
	% specific	55.0	23.0	53.4	58.0	
	2nd ½ % general	12.2		9.0	17.4	13.55
	% specific	87.8	13.8	91.0	82.6	
				N=20	N=9	
TASK III	1st ½ % general	45.7		44.8	47.0	
	% specific	54.3	20.25	55.2	53.0	
	2nd ½ % general	12.3		15.4	7.6	12.2
	% specific	87.6	15.0	85.6	92.4	
				N=16	N=12	

MEANS AND STANDARD DEVIATIONS FOR CONTEXT AND STATE ITEMS AS A FUNCTION OF TASK HALF AND ACCURACY

Table 2.2 presents the mean percentage scores of general and specific questions for the first and second half of the task. Standard deviations were also computed. The means of the percentage of general and specific question use for the accurate and inaccurate subjects is presented. The mean number of questions asked overall range from 12-13.55.

The tasks had been divided into two halves and the proportion of general and specific questions were calculated for each half. This allowed an exploratory analysis of the data.

TABLE 2.3

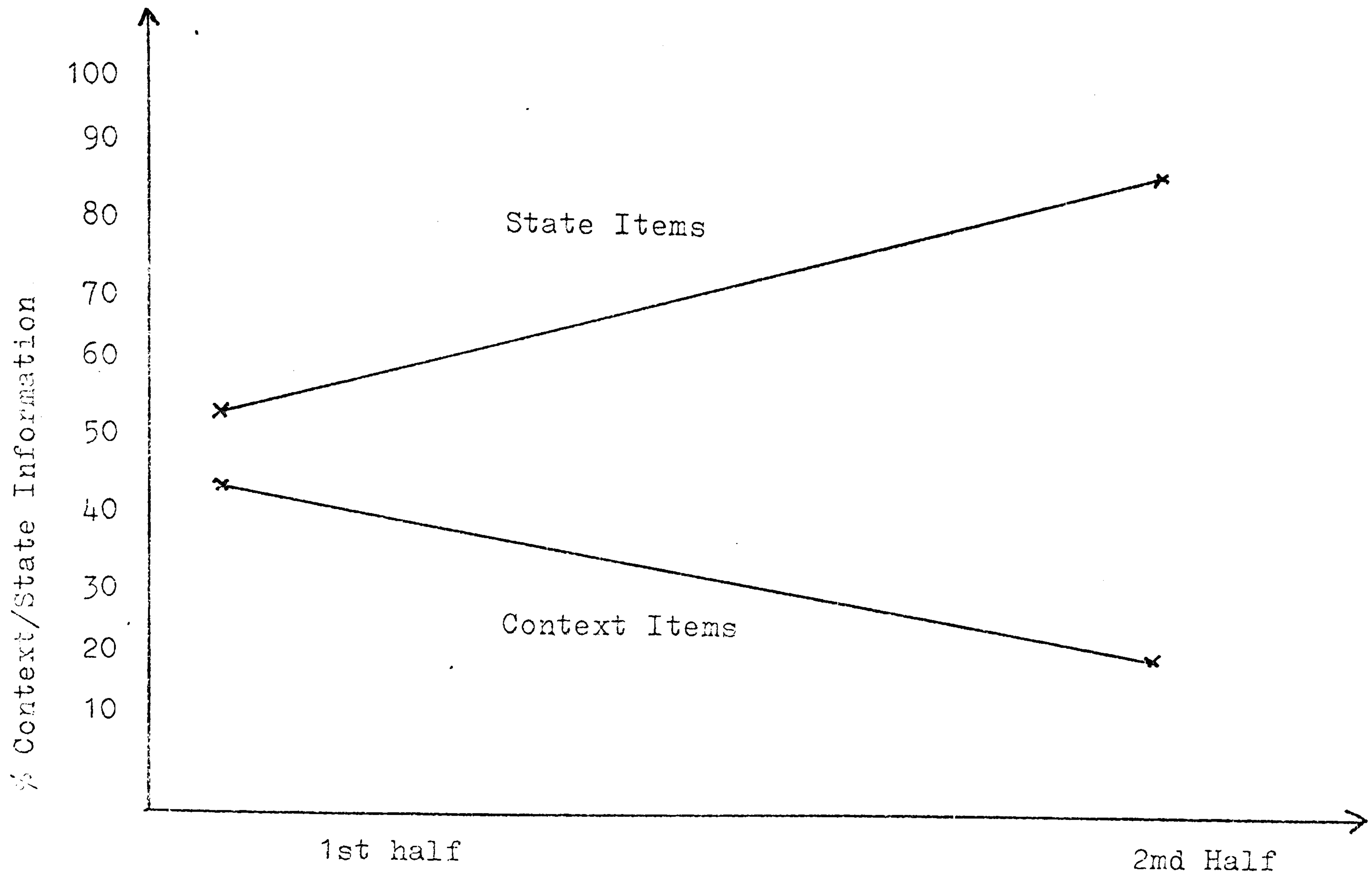
ANALYSIS OF VARIANCE SUMMARY TABLE WHERE GENERAL INFORMATION IS A FUNCTION OF TASK HALF AND COGNITIVE COMPLEXITY - TASK 2

SOURCE VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	P<
MAIN EFFECTS	15337.8	2	7668.9	21.14	.001
COMPLEXITY	58.8	1	58.8	.162	.689 (n.s)
TASK	15279.02	1	15279.02	42.13	.001
COMP x TASK	425.703	1	425.7	1.2	.284
EXPLAINED	15763.556	3	5254.519	14.490	.001
RESIDUAL	18857.213	52	362.639		

An analysis of variance was performed on Task 2 where context information items were examined as a function of task half and cognitive complexity (Table 3), task half being either the first or second half of the task. The Anova performed was from the S.P.S.S statistical package, the assumption of independence of the measures employed was violated in this design. The scores for the state and context items were correlated, each set of scores being percentages totalling 100%. In the main experiment it will be necessary to analyse the data in a manner that does not violate assumptions. However, this was not considered to be of major importance at the pilot stage.

The main effect of a significant difference between the use of context information as opposed to state information items was present ($F = 21.14$: $p < 0.001$).

Context and State Item use as a Function of Task Half - Task 2



The main effect for half of task was also highly significant ($F = 42.12$: $p < 0.001$). Graph I illustrates that this finding is a reflection of the use of more context items in the first half of the task and more state items in the second half.

There was no main effect of cognitive complexity, nor was there an interaction of complexity with any of the other variables.

Anovas of practise Task 1 and post-intervention Task 2 showed a similar pattern:-

MAIN EFFECTS

	CONTEXT	TASK ½
TASK 1	$F = 12.9$ ($p < 0.001$)	$F = 25.7$ ($p < 0.001$)
TASK 3	$F = 25.18$ ($p < 0.001$)	$F = 49.096$ ($p < 0.001$)

The post-intervention task did indicate a clearer distinction between the task halves and the use of the two types of items. The use of context information also increase compared with Task 1, although this difference in F value was considerably less between tasks 2 and 3 ($F = 21.14$ $F = 25.71$).

Separate Anovas looking at the state item use for the three tasks were not separately computed. The state and context scores being perfectly correlated, totalling 100, this Anova would be a mirror image of that examining the context item use.

Tests for homogeneity of variances for the three tasks were computed:-

Task 1 Variable:- Context information.
Cochran's C (13.4) = .38913, $p < .178$
Bartlett-Box F (3.4867) = 1.31917, $p < .267$

Task 2 Variable:- Context information
Cochran's C (13.4) = .4392, $p < .06$
Bartlett-Box F (3.4867) = 2.43743, $p < .063$

TASK 3 Variable:- Context information
Cochran's C (13.4) = .32337, $p < .564$
Bartlett-Box F (3.4867) = .70157, $p < .551$

The Null hypothesis was, therefore, not rejected at the p 0.01 level, and homogeneity of variance can be assumed for each of the tasks.

The data for subjects who were correct compared with those who were incorrect (see Table 2.2) suggests a trend in the direction of a reduction of use of context items by the more accurate subjects in the second half of the task for Tasks 1 and 2, the reverse being the case for Task 3. Students t-tests for independent groups were performed and no comparison reached an acceptable level of significance (App. 10).

Comparison was made between subjects who asked the general questions:-

7. Type of surgery
8. Time since surgery
39. Disease state

earlier in the task in terms of the mean rank order as a function of accuracy, and this is displayed in Table 3.

TABLE 2.4

	TASKS 1	2	3
Correct	9.5 (N=17)	5.2 (N=20)	4.2 (N=17)
Incorrect	10.66 (N=12)	9.7 (N=9)	3.7 (N=12)

MEAN ORDER OF USE OF ITEMS 7, 8 and 39

This comparison was not subjected to statistical analysis. A trend is present for the more accurate subjects in Tasks 1 and 2 to use these items earlier in the task. However, this trend is reversed in the third task. More subjects were incorrect in the third task and this could be an effect of confounding with task difficulty which will be discussed in the following section.

TABLE 2.5

TASK ACCURACY BY DIAGNOSTIC CONDITION

	A. PARALYTIC ILEUS (corr)	(incorr)	B. PULMONARY EMBOLISM (corr)	(incorr)	C. URINARY RETENTION (corr)	(incorr)
TASK 1	10	7	2	3	5	2
TASK 2	10	2	4	6	6	1
TASK 3			5	9	12	3

n = 29

This table reflects an effect of practice for the diagnoses paralytic ileus and urinary retention, where performance overall is more accurate in the 2nd or 3rd task than in the 1st task. A similar trend is not apparent for the condition pulmonary embolism.

It is also apparent that the balancing procedure for the diagnostic tasks is far from perfect. Both these observations will be discussed at length in the following section, as any result must be interpreted in the light of them.

Discussion

The results of this pilot study demonstrated a highly significant differential use of general and specific questions between the 1st and 2nd halves of the test ($p < 0.001$). General information served to start processing and lead to the use of specific information about the condition of the patient.

However, no further interpretation was appropriate because of an unknown effect of task difficulty. The bulk of errors in task 3 occurred in the diagnostic condition, pulmonary embolism. This was the only condition of the three that did not improve with practise. It was partly compounded by the partial balancing procedure which necessitated condition A, paralytic ileus, appearing in tasks 1 & 2, and conditions B & C being over-represented in task 2 & 3. This shortcoming in the balancing would have been justifiable had the tasks been perfectly equal in difficulty, which they patently were not. The reason for partial balance as opposed a fully balanced design was because the condition A was the example in the intervention.

The intervention, in its present form appeared to have been of doubtful effect. Although there was no formal statistical analysis Table 2.5 illustrates for the diagnosis pulmonary embolism there was a deterioration in performance across the tasks, more people were wrong in task 3 than in task 1 (9 v 3). The same pattern to a much lesser extent was observed for the diagnosis, urinary retention (3 v 2). But the confounding of task difficulty for the diagnosis pulmonary embolism made definite interpretation equivocal. It was therefore decided to retain the intervention in a modified form dropping the example of the diagnosis paralytic ileus. It was also considered to be advisable to concentrate solely on the role of general and specific information in memory and drop the example of stereotype as information sources - that may merely have served to confuse people.

Inspection of the list of complications revealed that the reason for the difficulty with task B was possibly due to an over-representation of respiratory conditions. Apart from pulmonary embolism, there were three alternatives. Chest infection, atelectasis and asphyxia were listed as possible diagnoses. But, there was only one complication of the urinary tract and one affecting gastric motility. It seemed sensible to equate the three systems for probability of occurrence.

Although the measure of cognitive complexity showed no association with the type of questions asked it was decided to retain this measure because of analysis of a larger sample and extreme scores could have been of interest.

CHAPTER 3

Experiment 1

The research hypotheses tested were:-

1. A higher proportion of general questions would be asked in the first half of the task in order to set the patient context compared with the second half of the task.
2. People who were relatively less cognitively complex would ask more general questions.
3. A decrease in accuracy would be associated with the use of general questions in the second half of the tasks.
4. People who asked their general questions earlier in the task would ask fewer questions overall than those asking them later because they reduced the number of available diagnostic hypotheses that needed to be tested.
5. This strategy was considered to be a more efficient search strategy and could be associated with increased accuracy.

3.1 Method

Design

The independent variables were:-

1. Type of information available
    - a) General information
    - b) Specific information

2. Cognitive complexity

3. Modified teaching intervention (App. 11). The dependent variables were:-
    - a) Accuracy
    - b) The number of questions of the different types of information

General Information was again defined as relatively stable facts about the condition of a patient.

Specific Information was defined as changeable facts about the condition of a patient.

The task was to diagnose the current condition of a general surgical patient.

Again, the only information given was that the patient had received general surgery (the general surgical condition was again appendicectomy).

The three post-operative complications for clinical diagnosis were:-

A = Paralytic ileus
B = Pulmonary Embolism
C = Urinary retention

The sequence of presentation was fully balanced and subjects were allocated in serial progression to the different orderings.

Patient Conditions	Subjects
A,B,C	8
A,C,B	8
B,A,C	8
C,A,B	8
B,C,A	8
C,B,A	8

The first task provided practice prior to the second task, the intervention and the third task.

Questions were again asked sequentially and on the basis of recognition of an appropriate question from the list (App. 3).

Subjects

Forty-eight student nurses in their third year of training at a London Teaching Hospital. The age range was 20-35 years and the range of educational achievement was GCE 'O' level to degree level, the majority of subjects being educated to GCE 'A' level. This represented two sets of student nurses who had completed their surgical module and had all had surgical nursing experience.

Materials

1. A two-part questionnaire (App. 2) assessed:-
   a) Personal details, ie. age, sex and educational qualifications
   b) Cognitive complexity measure
2. The standardised lists of questions for requesting information, as before (App. 3).

3. An extended list of 22 possible patient conditions, including the three target conditions, in randomised order (App. 12).

4. Pack of 39 individual "playing" cards with the answer to a request on one side and the sequence in the list of questions on the reverse. There were three packs, one for each diagnostic condition.

5. Modified teaching intervention dropping both the example of paralytic ileus and the discussion of stereotypes (App. 11).

6. The task instructions.

Procedure

The procedure was the same as that of the pilot studies. A general introduction and instructions were given initially (App. 6) followed by the completion of the questionnaire. The task instructions were read by the subjects at the same time as the researcher read them aloud (App. 7) modified for the inclusion of specific instructions for completing the tasks in silence and with no collaboration. Queries were answered by the researcher on an individual basis. The first two tasks were completed, they were followed by the intervention and the third task.

3.2 Analysis

An analysis of variance was computed to test hypotheses 1 and 2, and t- tests were computed to test hypothesis 3. Further ANOVA's were computed to test hypotheses 4 and 5 more directly, looking at the position in the tasks that the general questions were asked in relation to the total questions asked and accuracy. The results section will give details of how subjects were divided and how the measures were derived.

3.3

Results

To test hypotheses 1 and 2 an analysis of variance for one between and two within subject variables was computed. The anova was for the special case of a nested 3 factor design with repeated measures (Winer 1962: Appendix 13). The between subjects variable was cognitive complexity ie. subjects were divided into two groups of 16 representing the highest scoring $\frac{1}{3}$ on the Rep Grid, and the lowest scoring $\frac{1}{3}$. This represented subjects who were assumed to be relatively less complex and those who were relatively more complex, respectively. The within subject variables were:-

1. The three tasks that were sequentially completed. This was regardless of the diagnostic conditions which were completely balanced overall.

2. Each task was divided into two halves and the first half was then compared with the second. The comparison was made on the basis of the % of general questions asked of all questions asked in the first half of the task in relation to the second half. That is the % of general questions and the % of remaining questions were calculated for both halves of the task. The % of general questions were then compared.

A Hartley test for homogeneity of variance was computed (Appendix 13), where F max was observed as 2.2, d.f. (6,31). From tables at the level of $p<0.01$, d.f. (6,30) $F = 3.81$. Therefore the observed value fails to reach an acceptable level of significance and the Null Hypothesis need not be rejected.

Table 3.1 gives the ANOVA summary table (App. 13).

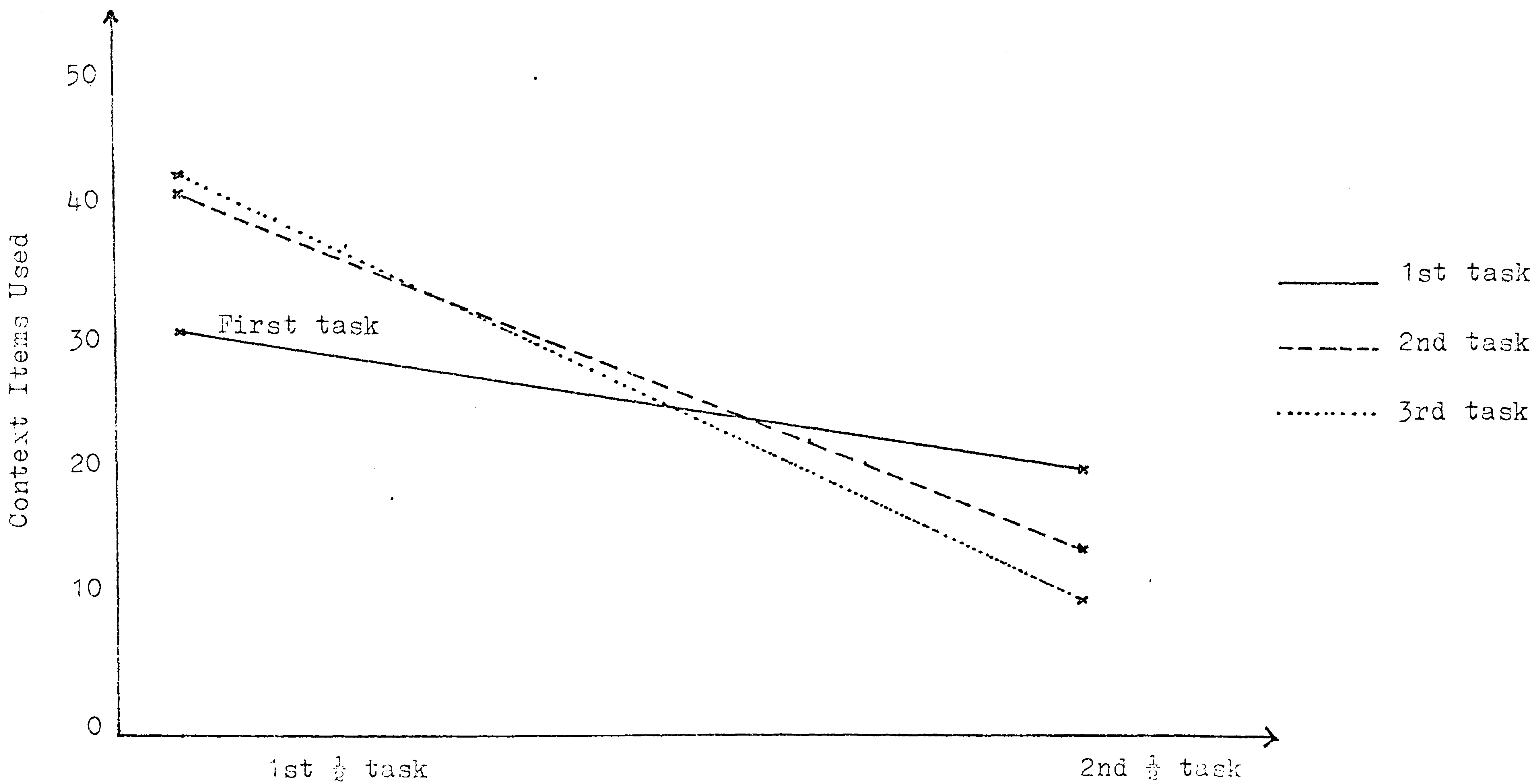
Table 3.1

ANOVA of Cognitive Complexity, the 3 Tasks & Half of Task

Source of Variance	d.f.	Sums of Squares	Mean Square	V.R.	P
Cognitive Complexity (C.C)	15,30	3.5	0.233	.0006	ns
Tasks	2,60	186.6	93.3	.78	ns
C.C. x Tasks	2,60	379.4	189.7	1.6	ns
Half of Tasks	1,30	26273.5	26273.5	142	<.001
Task x Half of Task	2,60	4270.24	2135.12	4.9	<.025
C.C. x Tasks x Half	2,60	155.06	77.5	.177	ns

A main effect of half of task was observed ($p < 0.001$) where significantly more general questions were asked in the first half of the task compared with the second half, this is shown in Graph 3.1. There is also a task x half of task interaction ($p < 0.025$), where the differences in the use of general questions between the two task halves increases over tasks (graph 3.1). No other observations reached an acceptable level of significance. Levels of Cognitive Complexity showed no differential use of information.

Mean Percentage of General Information Used as a Function of Task Half



The third hypothesis that people who ask relatively more general questions in the second half of the task compared with other subjects are also less accurate than other subjects was tested. T-tests for independent means comparing general information use in the second halves of the tasks between accurate and inaccurate subjects were calculated for tasks 1,2 & 3. No significant differences were observed (App. 14,15 & 16).

Hypotheses 4 and 5 suggested that an efficient strategy would result in fewer total questions asked and increased accuracy. Table 3.2 shows the overall means of the total number of questions asked by correct and incorrect subjects.

Table 3.2

Means of the Number of Questions Asked as a Function of Accuracy

	Correct	Incorrect
Task 1	15.4 (33)*	16.4 (15)
Task 2	13.5 (27)	13.9 (21)
Task 3	13.4 (39)	15.9 (9)

* Number of subjects

There was a very slight tendency for accurate subjects to ask between •4-2.5 fewer questions, this would not be a significant effect.

An analysis of increasing accuracy over the tasks was not performed. There was no simple increase in accuracy, rather a large increase between tasks 2 and 3 ie. 21 subjects were wrong in task 2 and only 9 in task 3. However this can be seen to be offset by a decrease between tasks 1 and 2 ie. 15 cf. 21.

In order to look at hypotheses 4 and 5 in more detail, the general questions that could be used to set the patient in a context were looked at specifically. These general questions were:-

6.	Type of admission	26.	Age
7.	Type of surgery	27.	Gender
8.	Time since surgery	28.	Weight
9.	Time in theatre	39.	Reason for surgery/disease state

Table 3.4 shows the number of general questions asked as a whole by accurate and inaccurate subjects.

Table 3.4

Number of General Questions Asked in the Tasks as a Function of Accuracy

	Correct	Incorrect
Task 1	3.33	3.7
Task 2	3.3	3.9
Task 3	3.3	3.2

There can be seen to be no appreciable differences in the numbers of general questions asked by accurate and inaccurate subjects, nor is there any difference between the tasks.

Subjects were divided into three groups for each task. The EARLY group asked the general questions that they asked more often in positions 1-3 than later in the task. The EQUAL group asked the general questions that they asked as often in positions 1-3 as later in the task and the LATE group asked their general questions later in the task. Table 3.5 shows the means for the three groups and tasks.

Table 3.5

Table of the Means of the Total Numbers of Questions Asked for the Three Groups

More General Questions		Task 1	Task 2	Task 3
Correct Subjects	EARLY	12.8 (14)*	12.2 (17)	12.9 (27)
	EQUAL	14.1 (7)	13.8 (5)	12.4 (9)
	LATE	19 (11)	16.25(8)	20 (3)
Incorrect	EARLY	14.75(4)	14.2 (12)	16.25(8)
	EQUAL	14 (3)	11.25(4)	18 (1)
	LATE	18.25(8)	19.25(2)	

* n

Frequently distributions were graphed out for tasks 1 and 2 (App. 17 & 18). Visual inspection suggested distributions within normal ranges, where in task 1 the mean was 15.6 and the median 14, and in task 2 the mean was 13.8 and the median was 14.

Two 2 x 2 ANOVAs for unequal cell frequencies with a least-squares solution (Winer 1962) were computed for tasks 1 and 2. The tasks were not combined into a single ANOVA because the derived measures would then have been correlated, correct and incorrect subjects could change categories between tasks. No analysis was made for task 3 because of a markedly skewed distribution in the direction of the early group.

The analyses considered the position that the general questions were asked enabling subjects to set the patient in context, in relation to the total number of questions that they asked. In task 1 a main effect of questions asked demonstrated that subjects asking the general questions that they asked earlier in the task asked significantly fewer questions overall ($F=6.4$, d.f. $(2,41)$ $p<0.01$: App. 19). The interaction between accuracy x questions asked failed to approach an acceptable level of significance ($F=.26$, d.f. $(2,41)$ $p<.25$).

The analysis for task 2 failed to reach the 5% level of significance, however a trend was observed in the same direction as in task 1 ($F=3.03$, d.f. $(2,42)$ $p<0.10$: where $F=3.23$, $p<0.01$: App. 20).

An alternative method of analysis looked at frequencies. Table 3.5 illustrates the number of correct subjects asking general questions early in the tasks 1, 2 and 3 to be 14, 17 and 27 respectively, compared with 11, 8 and 3 later in the task. A X^2 analysis found this to be a significant difference at the $p<0.025$ level ($X^2 = 8.175$, d.f. (2) $p<0.025$: App. 21). Correct subjects who asked their general questions equally often early or late were dropped from the analysis because the expected frequencies differed with respect to the observed frequencies in terms of less than one whole number.

A similar trend was observed for inaccurate subjects. Fisher exact probabilities were computed because of the small cell values involved. This necessitated 2, 2 x 2 matrices rather than 3 x 2, therefore task 1 was first compared with task 2 and then with task 3. The comparison between tasks 1 and 2 revealed no significant differences, However the comparison between tasks 1 and 3 showed that as the tasks progressed subjects asked the general questions at the beginning of the task. In task 3 no subject asked general questions later in the task. This effect was significant at $p<0.025$ (App. 22 and 23).

3.4

Discussion

The results offer some support for the experimental hypothesis of an initial strategy of setting a patient context that could serve to direct subsequent diagnostic hypotheses. In the first task that subjects performed the people who asked useful general questions initially asked fewer questions overall compared with the people who asked these same questions at a later point. However, there can be seen to be no clear relationship between this strategy and diagnostic accuracy within this experiment, which runs contrary to the findings of Gordon (1980). Gordon's work suggested that subjects who were less efficient in reducing the number of available alternative diagnoses were also less accurate as a direct result of this inefficiency.

Taking the overall use of general questions, both those that could be deemed to be useful in reducing alternative diagnoses and those that would have been less useful, subjects did ask more of these questions in the first half of the task than in the second half ($p < 0.001$). This effect increased as a result of practise as the tasks proceded ($p < 0.025$). The conclusion of the increased overall use of general questions earlier being a practice effect and not due to instructions in the teaching intervention, can be seen from inspection of Graph 1. The major gain for general question use in the first half of the task is between tasks 1 and 2, as is the major loss of these items from the second half.

Differential levels of cognitive complexity failed to predict the use of general as opposed to more specific questions by subjects. A hypothesis that a relatively less complex person could be predisposed to ask more general questions initially and fewer questions overall was not supported. The hypothesis was tenuous and the results observed question whether constructs could be seen to be synonymous with facts about a person. It was suggested that this was a poor piece of theorising and was subsequently dropped from the design.

The conclusion that increased accuracy between tasks 1 and 3 was the result of practice and not the result of teaching people to ask general questions initially was supported in several ways. There was no significant differences in any of the tasks when the proportion of general question in the second halves of the task was compared between accurate and inaccurate subjects. The same number of useful questions were asked by both accurate and inaccurate subjects, and no significant differences in initial useful question use were observed. In task 3 both groups of subjects asked their useful questions earlier than they had done in tasks 1 and 2. This may well have been the result of the teaching instructions but was unrelated to accuracy. Therefore, the teaching intervention can be seen to have had no useful effect.

The more interesting observation was that in task 1 those people asking their useful questions earlier in the task did ask fewer questions overall ($p < 0.01$). However this effect was not observed in task 2. No analysis was performed for task 3 because of the paucity of subjects asking these general questions later in the task. A converging observation, however, was that a significantly increasing number of correct and incorrect subjects asked these questions more in positions 1-3 than later as the tasks progressed. The effect observed in task 1 suggests that by setting a patient context initially subjects were then able to reduce available diagnotic alternatives and arrive at a decision earlier than subjects who did not adopt this policy. These subjects could be right or wrong in this decision. Two factors are evidently present, one of early patient context leading to earlier foreclosure and another involving accuracy. Both of which improved as the result of practise with the tasks, but are apparently unrelated.

In order to suggest possible reasons for these observations it is necessary to consider design features. The design of this experiment differs in several ways from Gordon's study. The tasks involved recognition rather than recall and recognition decisions. Gordon asked her subjects to recall from memory the questions that they needed to ask in order to diagnose the current condition of the patient. The only information that was provided was that subjects were told that the patient had had general surgery and subjects were given lists of possible diagnoses. Two conditions were employed, a condition of unlimited information and a condition of information being limited to 12 questions. The present experiment employed only the unlimited condition, it also presented subjects with a list of 39 questions that they could ask.

Gordon found that where the same group of subjects were restricted to the use of 12 questions, their performance was vastly superior to when they could ask an unlimited number of questions. This observation was complicated, however, by a failure to balance the diagnoses between conditions although they were balanced for practice. In the unlimited condition inaccurate subjects were characterised by asking for general information later in the task. That the diagnostic tasks were not balanced does not explain the association in the unlimited condition between inaccuracy and the late use of general information.

This finding was not replicated in the present case and could be attributed to the use of a recognition paradigm which was confined to a limited set of 39 items. In a recall paradigm where people had a free hand the initial strategy of setting the patient in context and eliminating unlikely diagnoses may contribute to a more accurate decision. For subjects who failed to employ this strategy the task could have been very difficult. They had to do several things, they had to retain all the information that they had asked for in a form of rehersal loop, Baddeley & Hitch, (1974). When eventually requesting information that could help to reduce decision alternatives, the working component of memory would be involved in correctly recognising appropriate items from those already requested, as well as retrieving questions that they may wish to ask.

In the current case these effects would be avoided by the subject's record of questions already asked and answers obtained from a limited set of 39 items. Therefore, where subjects had failed to reduce the alternative diagnoses to manageable proportions they could be assured of success by correctly recognising an abnormal sign or symptom that had been either radomly accessed or had been asked using the list of complications as a guide. The complications did break-down into discrete areas ie. cardiovascular, pulmonary, urinary, abdominal and either they were infective or non-infective. The inaccurate subject can be seen to have been wrong because they were inexperienced and had insufficient knowledge. The design employed did allow several strategies to be efficient in order to succeed with the decision.

The finding that asking the useful general questions earlier in the tasks resulted in asking fewer items overall is worthy of further study. It may be the case that this influences accuracy when the total set is less manageable, directing the information search when the burden on memory is removed. It would be equally interesting if this is not the case because asking nurses to write on a pre-prepared check-list the information they need to know may be a helpful way to improve performance in the wards.

CHAPTER 4

Experiment 2

It was considered to have been premature to attempt to teach a specific strategy at a stage where the effectiveness of that strategy was equivocal. The intervention did not have any beneficial effect and was therefore dropped from the design. This experiment concentrated on the use of different types of information. The design was adapted to include a large pool of questions as well as the small pool. Also, in an attempt to simulate time pressure, a condition was employed where subjects were restricted to asking 12 questions rather than as many as they desired. The recognition paradigm was retained. It was argued that with a large pool of potential questions that offer a large number of alternatives, asking questions in a non-systematic fashion would be counterproductive. It was considered to be more difficult to simply ask questions randomly until a relevant symptom was recognised.

The Research Hypotheses were:-

1. People who set the patient in a general context early in the task will ask fewer questions overall than those who do not adopt this strategy. They may also be more accurate.

2. When subjects are confronted with a large pool of questions and yet are limited to asking 12, those asking a larger proportion of general questions in the first half of the task will be more accurate as a result of this strategy. A strategy that reduces the number of possible diagnoses that need testing and leaves the later part of the task free to seek the relevant information about the current condition of the patient.

3. When there is a large pool of questions and only 12 can be asked, those who ask a smaller proportion of general questions later in the task will be more accurate as a result of this strategy. This would increase the opportunity to assess the current state of the patient.

4.1

Method

Independent Variables:-

1. Pool size (a) small question pool of 39 potential questions
   (b) large question pool of 76 potential questions

2. Restriction applied (a) no restriction of the number of questions that could have been asked
   (b) restricted to ask 12 questions

3. Type of information available (a) general information
   (b) specific information

Dependent Variables:-

1. Accuracy

2. Total number of questions asked

General and specific were defined as before and the task requirement was identical, as was the information given. The operations selected were appendicectomy and lumbar sympathectomy. The two possible conditions to be diagnosed were:-

(a) urinary retention - (C)

(b) wound infection - (D)

The combination of diagnoses, restriction and pool size were fully balanced, as were subjects, over the tasks.

Conditions

Task I	C	UR	LP 1, 9, 17	81
		R	SP 3, 11, 19	83
		R	LP 7, 15, 23	87
		UR	SP 5, 13, 21	85
	D	R	LP 2, 10, 18	82
		UR	SP 4, 12, 20	84
		UR	LP 6, 14, 22	86
		R	SP 8, 16, 24	88

Task II	D	UR	LP 3, 11, 19	83
		R	SP 1, 9, 17	81
		R	LP 5, 13, 21	85
		UR	SP 7, 15, 23	87
	C	R	LP 4, 12, 20	84
		UR	SP 2, 10, 18	82
		UR	LP 8, 16, 24	88
		R	SP 6, 14, 22	86

R = Restricted

UR = Unrestricted

LP = Large Pool

SP = Small Pool

Two practice trials were completed prior to the experiment trial

Practice	A	UR	LP
	B	R	SP

A = Paralytic Ileus

B = Pulmonary Embolism

Subjects

88 student nurses in their third year of training. The majority of subjects had been educated to G.C.E. 'A' level. There were a small number of graduates, but level of education was not taken into account in the analysis because the numbers of graduates were small.

Materials

The materials were the same as for experiment 1 except:-

1. An extended list of 76questions, the order of which was fully randomised. They were typed with the same layout as the short list on three separate sheets stapled together. The list of questions had been subject to expert scrutiny (App. 24).

2. Additional packs of cards for the extended lists and additional packs, short and extended for the new diagnosis wound infection. Again the sign and symptom cluster associated a particular complication were taken from Fream (1978). All potential questions and answers were subject to expert scrutiny (App. 25).

3. Modified task instructions (App. 26).

Procedure

A series of five identical experimental sessions was conducted in the same lecture theatre that had previously been set up for the procedure on each occasion. Beside each subject number were placed, face down, the practice lists and packs with the pre-arranged order of the tasks and conditions recorded on the top of each pack and associated list. Each subject also had a list of the 21 post-operative patient conditions and a set of typed instructions.

Participation was voluntary and anonymous and instructions were read by each subject and accompanied aloud by the researcher. The task was to judge the current condition of general surgical post-operative patient. The nurses were told that initially they would do this twice and that on each occasion the patient's condition was different. Controls for silence and collaboration were again instituted. The nurses were then asked to look at their lists of questions and they were told that they had a long list and a short list. The procedure of starting to ask questions by choosing a question anywhere on the list and then proceeding to ask subsequent questions, recording both the answers and the order asked was explained. It was explained that on one occasion they were limited to asking twelve questions and on the other occasion they could ask as many as they wished. These instructions were also recorded at the head of the question lists. The nurses started with their first pack and when they had arrived at and recorded a diagnosis proceeded with the second pack.

When the practice trials were completed the researcher gave the experimental material to each subject individually explaining the completion order of the two tasks, which task had a small and which had a large pool, and in which task they were restricted to asking 12 questions. This information was also stated on the materials. All subjects completed the experimental tasks. It was, however, noted by one subject that a rogue card was present in the large pool, diagnosis C, urinary retention. The card bearing the patient's current temperature stated it as 34[7] rather than 37, causing this subject to diagnose hypothermia. This discovery was made in the final experimental session. No other subject was misle d by the mistake but the misle d subject was excluded from the error analysis. The sessions took one hour to complete and on completion the materials were collected and a full explanation was given.

4.2

Analysis

The total number of questions asked were subjected to an analysis of variance to examine the effect of restriction, pool size and the two different diagnoses. The first hypothesis was tested by analysis of variance of the unrestricted data. T-tests comparing the % proportion of general question use in both task halves were computed to test hypotheses 2 and 3.

4.3

Results

Table 4.1

Means of Total Number of Questions Asked in Tasks 1 and 2 Under the Diagnostic Conditions C (Urinary Retention) and D (Wound Infection), with a Small/Large Question Pool and Restricted or Unrestricted Information

	Diagnosis C				Diagnosis D			
	Small Restr.	Small Unrestr.	Large Restr.	Large Unrestr.	Small Restr.	Small Unrestr.	Large Restr.	Large Unrestr.
Task 1	11	14.4	10.8	14	11	11.3	10.4	10.7
Task 2	10.6	13.45	11.4	14.1	9.2	13.3	10.8	13.8

Grand Means - Task 1 = 11.7
Task 2 = 12.08

n = 11 per cell

The total question use under the different conditions was examined. Inspection of the means in Table 4.1 indicated that in the unrestricted condition in Task 1 between 1-4 more questions were asked in diagnosis C compared with diagnosis D. This was not the case in Task 2 where question use appears equivalent. An ANOVA computed for Task 1 demonstrated that the difference in the number of questions asked between the two diagnoses was significant ($F = 5.377$, d.f. 1,80 $p<0.023$). In Task 1 subjects who were not restricted did ask more questions overall ($F = 5.965$, d.f. 1,80 $p<0.017$). However, this effect was shown to be confined to diagnosis C by the diagnosis x restriction interaction ($F = 4.038$, d.f. 1,80 $p<0.048$). The main effect of differential question use between diagnoses C and D, plus the interaction in Task 1 can be shown by inspection of the means in Table 4.1 to be due to diagnosis D, wound infection. There can be seen to be less than a whole number difference between subjects under differing levels of restriction, suggesting that this was an easier diagnosis to make than C. The ANOVA computed for Task 2 showed no differences in question use between the diagnostic conditions and a significant main effect of restriction ($F = 15.44$, d.f. 1,80 $p< 0.001$), unrestricted subjects consistently asked more questions in both diagnostic conditions (App. 27).

To consider the possible effects of the rogue card in the large poole of questions in diagnosis C, the effect on subjects of starting to make this diagnosis first were considered.

Table 4.2

Difference Between the Means of the Total Questions Asked in Task 1 and 2 Where Diagnosis C (Urinary Retention) was the First Task and Diagnosis D (Wound Infection) was the Second

		1st Condition - Diagnosis C			
	2nd Condition - Diagnosis D	Small Unrestr.	Large * Unrestr.	Small Restr.	Large* Restr.
Small Restr.			4.8		
Large Restr.		3.6			
Small Unrestr.					2.5
Large Unrestr.				2.8	

Where subjects were exposed to the large pool, diagnosis C, the differences between the mean questions asked between the two tasks were very similar to those who were not exposed to this condition ie. 4.8 cf 3.6 and 2.5 cf 2.8.

Table 4.3

Difference Between the Means of the Total Questions Asked in Tasks 1 and 2 Where Diagnosis D (Wound Infection) was the First Task and Diagnosis C (urinary retention) was the Second Task

		1st Task - Diagnosis D			
		Small Unrestr.	Large Unrestr.	Small Restr.	Large Restr.
Small Restr.	2nd task - diagnosis C		.1		
Large Restr.		.1			
Small Unrestr.					3
Large Unrestr.				3.1	

Where subjects were exposed to the large pool, diagnosis C as the second task again the difference between tasks 1 and 2 were similar to those subjects who did not encounter this pool ie. .1 cf .1 and 3.1 cf .3. Encountering a rogue card initially in the first task did not appear to exert a dramatic effect over the nurses performances. It was, therefore, concluded that the diagnosis C, urinary retention, was more difficult than D, wound infection, but this effect was independent of the mistake with the temperature recording. Both the diagnoses had been fully balanced over conditions and tasks, therefore the diagnostic conditions were combined for further analysis.

Table 4.4 gives the overall level of subjects who were both accurate and inaccurate for both levels of restriction and size of pool.

Table 4.4

	Restricted		Unrestricted	
	Small Pool	Large Pool	Small Pool	Large Pool
Correct	33	31	31	36
Incorrect	11	13	13	7

one subject excluded as previously stated

In order to examine the hypotheses it was necessary to consider general question use.

In the small pool these were:-

6. Type of admission
7. Type of surgery
8. Time since surgery
9. Time in theatre
26. Age
27. Gender
28. Weight
39. Reason for surgery/disease state

In the large pool these were:-

2. Disease state
4. Frame size
6. Age
7. Alcohol consumed daily
20. Type of admission
29. Socio-economic status
31. Marital status

34. Time in recovery
35. Duration of anaesthesia
36. Cigarettes daily
39. Housing conditions
44. Wound closure technique
46. Past medical history
48. Pre-medication given
49. Pre-operative mobility
54. Type of anaesthesia
55. Weight
56. Occupation
59. Type of surgery
65. Gender
66. Past psychiatric history
67. Time since surgery
72. Height
73. Duration of surgery

All these questions were regarding the patient before or during surgery, questions that could be used to place that patient in a context, and thus influence expectations.

Subjects were then divided into two groups:-

1. Subjects who asked the general questions more often in positions 1-3 compared with later in the task - the EARLY group.

2. Subjects who asked their general questions either later in the task or as often late as early - LATER.

Means of the Total Numbers of Questions Asked by Early and Later Subjects as a Function of Pool Size, Restriction and Accuracy

Table 4.5

			Gen. Questions Asked EARLY	Gen. Questions Asked LATER
Restricted	Small Pool	correct	9.95 (20)	10.5 (13)
		incorr	11.3 (9)	11 (2)
	Large Pool	correct	10.8 (19)	10.25 (12)
		incorr	11.3 (7)	11.7 (6)
Unrestricted	Small Pool	correct	11.3 (25)	15.5 (6)
		incorr	14 (7)	17.2 (6)
	Large Pool	correct	12.5 (19)	14.5 (15)
		incorr	12.3 (3)	18 (3)

n in parenthesis

Inspection of Table 4.5 shows that, in the restricted conditions, very little difference between the pool sizes or accurate/inaccurate subjects was apparent. In the unrestricted condition there are greater differences between the means. In the unrestricted condition there was only approximately one question difference between the pool sizes and the appropriate sub-groups of early, later, correct, incorrect subjects. These were therefore combined to give a more meaningful sample size. Table 4.6 shows the mean of the total questions asked by early or later subjects in the unrestricted condition, both pools and who were either accurate or inaccurate.

Table 4.6

Unrestricted Condition; Combined Pools: Mean of Total Question Use as a Function Stage of General Question Use and Accuracy

	CORRECT	INCORRECT
EARLY	11.8 (44)	13.5 (10)
LATER	14.8 (21)	17.4 (9)

n in parenthesis

Visual inspection of the frequency distributions of the total questions used by the Early and Later groups revealed curves approximating to the normal distributions (App. 28). Therefore, an ANOVA, least squares solution for unequal cell frequencies and non-orthogonal date (Winer 1962) was computed (App. 29).

There was a main effect of stage of use of the general questions ($F = 9.69$, df $(1,80)$ $p<0.01$). People who asked their general question earlier asked fewer questions overall. There was a trend for inaccurate subjects to ask more questions overall ($F = 3.3$. df. $(1,80)$ $p<0.10$). The interaction accuracy x stage of asking general questions, was not significant. Considering the frequency data a X^2 was computed to compare the number of subjects asking general questions Early and those asking them later as a function of accuracy. This failed to approach significance (App. 30).

To test hypotheses 2 and 3 effectively general question use was considered in terms of the proportion of these questions asked in the first and second halves of the task ie. the proportion of general question use compared with more specific questions use was calculated for both task halves. This gave an arithmetic range 0-50%.

Table 4.7

Mean General Question Use in all Conditions - % Proportion

		Restricted		Unrestricted	
		Small Pool	Large Pool	Small Pool	Large Pool
Acc.	1st ½	19.9, n=33 (SD=9.1)	20.8, n=31 (SD=10.6)	20.8, n=31 (SD=10.6)	18.07, n=36 (SD=11.9)
	2nd ½	5.2, n=33 (SD=7.4)	2.9, n=31 (SD=5.35)	6.5, n=31 (SD=6.4)	6.5, n=36 (SD=7.02
Inacc.	1st ½	18.7, n=11 (SD=7.1)	16.9, n=13 (SD=9.5)	23.7, n=13 (SD=9.35)	12.35, n=7 (SD=6.7)
	2nd ½	4.2, n=11 (SD=9.2)	7.15, n=13 (SD=6.7)	6.9, n=13 (SD=10.1)	9.8, n=7 (SD=10.5)

To examine the critical condition of a large pool of potential questions where the subjects were restricted to asking 12 questions, the randomised test for two independent samples was considered to be a suitable test. The comparisons made were the percentages of general questions asked in the two halves of the tasks by accurate and inaccurate subjects. Where there are larger samples (n = 44) the T-statistic is recommended rather than the randomised test (Pitman, 1937: frequency distribution App. 31).

The t statistic was computed comparing the use of general questions by accurate and inaccurate subjects in the second half of the task. A larger proportion of the second half of the task was devoted to asking general questions by inaccurate subjects compared with accurate subjects ($t = 2.23$, df. 42 $p < 0.025$ for a 1-sided test: $p < 0.05$ for a 2-sided test, App. 32).

A comparison of general information use in the first half of the task failed to reach significance (App. 33). As did both comparisons of differential use in the unrestricted large pool condition and no trends were observed (App. 34 and 35). Graphs of the distributions are appended (App. 36-38). It is evident from an inspection of the means for both the restricted and unrestricted conditions, with a small pool of questions, that there is little variability in the use of general information by accurate subjects, compared with inaccurate subjects. This is graphically presented in App. 37 and 38.

4.3

Discussion

The rogue card present in the large pool of questions where the diagnosis was urinary retention will be discussed initially. In task 1, significantly more questions were asked by subjects making this diagnosis compared with those diagnosing a wound infection. The question was as to whether or not this was due to uncertainty induced in the subjects by an abnormal sign that was irrelevant. This does not appear to have been the case. Table 4.2 illustrates the difference in the means of the total questions asked where the diagnosis for the first task performed was urinary retention and the second was wound infection. The different scores observed where the first diagnosis was urinary retention compared with when this was not the case were 4.8 cf 3.6, and 2.5 cf 2.8. A difference of 1 between the observed means would not be statistically significant. It was therefore concluded that further analysis was not invalidated.

Hypothesis 1 received partial support. The people who did ask the general questions earlier in the task, which implied a deductive strategy of setting the patient in a general context at the outset, asked fewer questions overall. But this was not associated with increased accuracy. This was a replication of the findings of the first experiment and suggested that the nurses who asked their general questions early did adopt a strategy that served to reduce the number of available hypotheses initially. However, it appeared to increase their speed to decision rather than their accuracy. There was a trend for inaccurate subjects to ask more questions overall ($F = 3.3$, df. $(1,80)$ $p < 0.10$). But this overall asking of more questions was not associated with any particular type of question order. The trend could indicate a lower level of confidence amongst inaccurate subjects, but as this observation failed to reach the accepted 5% level of significance, it will not be discussed further.

Hypothesis 2 was not supported. The suggestion was that when time was limited and people were restricted to asking up to 12 questions from a large pool of potential questions, asking general questions early in the task would prove to be a more efficient strategy. The prediction was that a deductive and, therefore, directed search for information from the beginning would prove to be more fruitful. This was not demonstrated, accurate subjects did not ask a higher proportion of general questions in the first half of the task compared with their inaccurate colleagues.

Hypothesis 3 was supported. The inaccurate subjects asked a higher proportion of general questions in the second half of the task, when they had a large pool of potential questions and were restricted to asking 12. A general, context question asked later could have provided useful information to corroborate a diagnosis. But where there was a limit to the number of questions that could have been asked, the more that were devoted to obtaining background information proportionately reduced the opportunity to assess the current condition of that patient. If the subjects' objective had been to generate a further hypothesis, having rejected the first they then had insufficient time to test it and the response observed could have been a 'bad guess'. This was a similar finding to that of Gordon 1980.

Overall, the result demonstrated that an early use of general information providing a context for a patient did result in asking fewer questions in total ie. earlier foreclosure. This was a direct replication of the same effect present in experiment 1. The implication that people who adopted a particular search strategy were more decisive as a result of this strategy is of interest, and will be discussed further at a later stage in the discussion. But this would have been more interesting had it aided performance in the restricted case. From the results of this experiment there was only a slight indication that a search strategy adopted had a beneficial effect on the selectivity within attention, in terms of accuracy.

The evidence suggested that any beneficial effect of initial selection between hypotheses and subsequent testing in a paradigm that involved recall of what it was necessary to know and retention of that information, must have been associated with a reduction of the cognitive load. But, before stating definitively that there was no independent effect of search an inspection of the means is warranted. An inspection of the means from Table 4.7 reveals that there was 3.9 difference between the proportion of general information used in the first half of the task by accurate and inaccurate subjects and 4.25 in the second half. Yet, the difference in the second half proved to be statistically significant. This would be accounted for by the relatively small ratio of the difference between the means and the observed means in the first half compared with the second half (ie. 20.8 - 16.9 = 3.9 & 7.15 - 2.9 = 4.25). Had more subjects been incorrect the difference between the means would have been inflated. This is suggested further when inspecting Table 4.6, 50% of the subjects who asked general questions later were incorrect compared with 25% of subjects in the early group. The lack of significance of the X^2 could again have been a reflection of the tasks that were too easy and insufficient people were incorrect.

In the light of the significant observation in the second half of the task, coupled with the low number of inaccurate subjects, it is suggested that it cannot be unequivocally stated that there is no independent effect of a search strategy. It can be argued that an effect of cueing was still present with a provided list, however long. In terms of making the task more realistic and more like the situation encountered on the wards, information was provided and the decision required was that of recognition of appropriate items. But, the selection of information available is not as overt in a ward as it is with the provision of a list of potential questions. In order to conclude unequivocally that deductive reasoning does not organise search, ie. the selectivity within attention, in an orderly fashion that results in advantage for the nurse a further experiment is required. A further experiment employing a recall and recognition paradigm, where the nurses would be required to recall what they wish to know and record the answer to their questions.

Returning to the observation that an early use of general questions resulted in fewer questions being asked overall, this was found to be a robust observation. The suggestion is that deducing from a reduced selection of alternatives determines a finite number of pieces of specific information that need to be examined. Whereas, induction from specific information does not circumscribe the questions that are potentially useful. The repeated observation that early use of general information leads to earlier foreclosure, strongly suggested that these subjects were effectively reducing the number of alternative hypotheses in a deductive fashion. This reduction of alternatives was independent of a retentive load on memory.

The restricted number of alternative hypotheses were selected on the basis of their "representativeness" in terms of the initial general information. This lead to the examination or retrieval of a limited set of signs and symptoms associated with chosen hypotheses. This would be predicted by a script theory of memory where diagnostic scripts have a form of unique identity (Abelson, 1976). The Prospect Theory of Tversky & Kahneman (1981) predicts an initial editing phase based on the context of the problem. The context would trigger a predisposition to behave in a way determined by the perception of the task and allows accommodation of task specific effects. Prospect Theory would predict the early foreclosure in terms of cost-benefit, but the Script Theory of Abelson gives no information about mechanisms of more complex analytic decision processes.

4.4

Concluding Discussion

The objective of this part of the study being reported was to examine what could usefully be applied. Other work has demonstrated the strategy of an initial selection of hypotheses amongst doctors and nurses (Kassirer & Gorry 1978; Tanner 1977). It was not the intention of the study to add directly and exclusively to the theoretical understanding of decision making.

The objective had been to identify an aid to natural decision making that could be taught to inexperienced nurses. It had become evident that even with an improved experimental design, if there were any effects that had previously been obscured and could have been demonstrated, they would have a very limited applied value. The strength of an effect of increased accuracy due to adopting a particular strategy would need to be strong to effectively improve performance after a general teaching intervention.

The teaching intervention and the notion of attempting a teaching intervention, can be seen to have been naive in terms of attempting to teach something that had not been clearly demonstrated in the past. Although the possibility exists for a use of such 'games' by student nurses, they could help them to define the information they require when making a decision about the condition of a patient in a variety of situations.

CHAPTER 5

Literature Background for the Study of Nurses' Perceptions of Patient Affect

The second clinical judgement examined was nurses' perceptions of patient affect. Hospital patients are known to experience feelings of negative affect and depression, to a greater or lesser extent. Their feelings were examined in relation to the degree to which nurses were aware of their presence. Higgins (1980) observed that the main emphasis in the study of clinical judgements had been directed at the differential use of specified signs and symptoms. The interpersonal aspects of clinical judgement have received considerably less attention.

This half of the work reported was concerned specifically with nurses' awareness of patient feelings of anxiety and depression compared with physical feelings such as tiredness.

At the start of this review a brief account is given of work that identified the presence and significance of negative affect on hospital patients, thus presenting a rationale for the selection of this clinical judgement. In order to make any clinical judgement it is necessary to have information on which to base it. The potential information sources available to the nurse are considered. The potential sources that would allow a nurse to make a judgement about a patient relate to the verbal and non-verbal behaviour of that patient, as well as information based on the previous experience of that nurse having encountered patients in that context in the past. The relative use of each of the above as a potential source of information will be explored.

5.1

Anxiety and Depression in Hospital Patients

Initially it is necessary to consider, briefly, the nature of anxiety and depression and whether or not patients do experience this whilst in hospital. Spielberger (1972) defined anxiety in terms of an emotional state that is palpable but transitory and is associated with feelings of tension and apprehension, plus heightened autonomic nervous system activity, which is often due to the threat of the unknown. This can be otherwise termed as a state of anxiety, where the trait or predisposition to experience anxiety determines the ease with which anxiety is experienced plus the intensity. A definition of depression included a low subjective mood, pessimistic and nihilistic attitudes, loss of spontaneity and specific negative signs (Beck, 1967). Becker (1974) had observed the overlap between anxiety and depression in component feelings and in clinical manifestations, and preferred to group anxiety and depression as 'affective states' without distinguishing the separate emotions. But they are suggested to be subject to differences in cognitive appraisal ie. anxiety being a mobilisation to cope and depression, passive despair.

That is a brief, and necessarily superficial, attempt to distinguish between the two emotions anxiety and depression which has taken no account of degree. Anxiety has been implied by Becker (1974) to be the positive aspect of the 'affective state' implying coping and control. Whereas, depression is suggestive of the abandonment of attempts to take control given experiences of events that are not the consequences of actions taken (Seligman 1978; Miller & Norman 1979; Miller 1979).

To turn to whether or not patients experience negative affect whilst in hospital, Wilson-Barnett (1977) investigated aspects of ward life that evoked negative feelings for many people. The study employed an interview technique and patients were asked how they felt about 60 items that described various aspects of hospitalisation eg. admission to hospital, having set visiting hours, being away from family. This study isolated seven items that had engendered responses that implied anxiety or depression. The items were those directly relating to illness and separation from the persons normal environment, and fear of the unknown. Further work demonstrated feelings of anxiety and depression amongst medical ward patients (Wilson-Barnett & Carrigy, 1978). 30% of the patients studied who had been in hospital for five days or longer reported having experienced 'strong emotional reactions'. 'Strong emotional reactions' were also associated with higher levels of trait anxiety on the E.P.I. (Eysenck & Eysenck, 1964).

Other work has reported similar findings. Younger patients, those with a higher trait anxiety score and those with a better chance of recovery felt more distressed about being in hospital (DeWolfe, Barrell & Cummings, 1966). Cartwright (1964) and Hugh Jones, Tanser & Whitby (1964) identified specific patient concerns exacerbated by hospital admission. The fear of pain, the unknown and destruction of body image have been identified as patient concerns (Carnevali, 1966).

Maguire, Julier, Hanson & Bancroft (1974) studied two hundred and thirty medical patients in two wards. They were given a standardised questionnaire and a psychiatric interview. 23% were considered to be psychiatrically ill and affective disorders were the commonest disorders encountered. The prevalence of depressive and anxiety symptoms was equivalent among different disease classifications with the exception of the psychosomatic group. Moffic & Paykel (1975) in the United States found a similar profile, but with a higher overall percentage of patients who were depressed. These studies are not directly comparable because different criteria were employed, a problem when interpreting many studies of this type. Moffic & Paykel (1975) suggested that the severity of illness was directly related to the degree of depression. But with another group of patients anxiety and depression have been associated with the less severely ill and those recovering (Schwab, Bialow, Brown & Holzer, 1967). The overall percentage of patients affected was 22%, consistent with other studies.

Depression is known to characterise a number of physical disorders as a result of biochemical abnormality (Lipowski, 1967; Castelnuovo-Tedescu, 1961). The presence of diagnosed 'clinical depression' has been specifically associated with a higher level of life events than for those who are not depressed. (Brown, 1972; Kendell, 1976). Similarly a group of myocardial infarction patients were found to have past histories that reported more divorce, lonliness, sleep disturbances and feelings of nervous stress and depression compared with a matched group of controls (Thiel, Parker & Bruce, 1973).

Kendall & Watson (1981) have reviewed the area of psychological preparation for surgery and provided convincing evidence for a relationship between raised levels of anxiety and recovery. Having acknowledged this as a problem for hospital patients, a considerable body of work has been directed towards reducing anxiety. Reducing the anxiety by means of an intervention offering information or coping strategies prior to an operation or procedure (Hayward, 1975; Wilson-Barnett, 1977; Ridgeway & Mathews 1982). Further, there has been work tracing the natural cause of anxiety before and after surgery (Johnstone 1980). The suggestion was that high levels of anxiety were experienced before admission to hospital, between admission and surgery and following surgery. They were not restricted to the immediate pre-operative period.

Overall, feelings of depression have been found in a sizeable proportion of hospital patients, as well as the better documented anxiety. The presence of negative affect in hospital patients, especially at high levels that are protracted, and its association with interrupted recovery was isolated as an important potential patient problem in this study. The question that emerged related to the accuracy of nurses' clinical judgements about their patients feelings, their knowledge of the range and extent of feelings that patients could experience. This was considered important because appropriate intervention to facilitate recovery is dependent on knowledge of the patient's psychological condition.

The potential sources from which a nurse can obtain this information are several. How patients say they feel, their non-verbal behaviour and the context in which the patient is placed ie. the severity of the illness, its duration and possible outcome. These potential sources of information will be considered in turn.

5.2

Verbal Behaviour

The extent to which nurses talk to their patients has been the subject of recent study (Macleod Clark, 1983). The frequency, duration and content of nurse/patient verbal exchanges were investigated in a sample of surgical ward patients. The sample of 56 hours of tape-recordings contained 310 nurse-patient conversations and the average conversation time was 1.71 minutes, and there were approximately 5.5 conversations per hour. These recordings had sampled two-hour periods over the entire space of day duty. The nurses had spoken with their patients relatively infrequently and for short durations. The contents of the majority of conversations were related to a task being performed by the nurse. More importantly, enquiries regarding how the patient felt, on the rare occasion that they occurred, were posed in the form of a closed question designed to elicit a desired response. This was a small study on a single population of nurses, but it was well controlled and the results are supported by previous work. Goddard (1953) showed that nurses spent about 9% of their time in conversation with patients. Others have studied the time spent talking to patients and all have found that this represented a small part of the total day (Manchester Study 1955; Adams & McLlwraith 1963; Altschull 1970; Wells 1975; Pepper 1977). Studies with cancer and gynaecological patients have also suggested that nurses tend to avoid discussion of feelings (Pepper 1977; Bond 1978).

5.3

Nonverbal Behaviour

The study of accuracy of social perception fills a large literature and divides into two broad areas. Firstly, there is the study of the recognition of emotion which centres largely around perception of facial expression of emotion. Secondly, there has been work, most of which pre-dates 1965, concerned with the accuracy of judgement of enduring dispositions that are assumed to be present, within other people. This work also assumed a generalised ability to judge others. The first area will be considered in some depths and the second in terms of methodological problems encountered in measuring accuracy. The methodological problems involved in studying the accuracy of subjects' perceptions of anothers nonverbal messages are numerous. One of the chief problems is that of the criteria adopted to gauge the success of the judgement.

The early work concerned with whether or not perceivers are accurate in judging emotional states on the basis of facial expression was initiated by Darwin (1872). He considered that facial expressions were innate rather than learned within a culture and universal within a particular species. The method that he employed was adopted for subsequent research. This was the 'judgement study'. The investigator showed pictures or photographs of facial expression to a number of subjects and asked them to judge the feelings that were being expressed. Darwin had concluded that people could judge emotion accurately, although there were several sources of inaccuracy. There was variability in the difficulty of the judgement task, between different photographs and different emotions. Schneider, Hastorff & Ellsworth (1979) point out that Darwin recognised that a judge's own emotional state could bias the interpretation of the state of another. They also note the influence of expectations that the perceiver holds or the context on perception and ask whether peoples' judgements are a result of these biases rather than there being any underlying ability to judge facial expression accurately.

The research question that dominated the early research in this area was the question of an innate link between emotion and facial expression. Alternatively, the facial expression of emotion could be determined within a culture and transmitted through learning. This has obvious implications for nurses nursing patients from other cultures. The early research cited in Schneider, Hastorf & Ellsworth (1979) concluded that not only was there no innate link but that there was little or no association between the face and emotion. However, the methodological problems that Schneider et al. enumerate make any interpretation of the data dubious. The judgement study designs employed inadequate stimulus materials and tended to study a very limited range of similar emotions, This made the judge's task to accurately discriminate an impossible one.

Other studies that attempted more naturalistic designs also had problems. They employed unposed subjects to be judged when placed in situations that were intended to illicit a particular emotion. They failed to control the variety of emotions that could be elicited in different people exposed to the same stimuli, plus their possible reluctance to express it under certain circumstances (Landis 1924;1929). Sherman (1927) concluded that people could not judge emotion accurately on the basis of their failure to discriminate the emotions he supposed two babies less than one week old to be experiencing. The result of the eliciting stimuli on all occasions was a crying baby.

The work that followed was based on the conclusion that people are unable to judge emotion from facial expression and studied the judge and environmental cues or context. However, Frijda (1969) observed that judges tended to rely on facial expression rather than context when the facial expression was unexpected within a given context. Mehrabian & Wiener (1967), Mehrabian & Ferris (1967) found that facial expression was the most important cue for subjects making judgements of emotion.

Consideration of the face in relation to other sources resulted in a spate of studies that took an information theory 'channels of processing' approach to research in this area. They dubbed the various information sources as 'channels' ie. the voice, speech, body movement and face and attempted to examine the relative importance of each channel for the perceiver. Most experimenters found that the face was more accurately judged, produced higher agreement between judges or correlated better with judgements based on full audio-visual input than did the voice or speech alone (Argyle, Alkema & Gilmour 1971; Burns & Beier 1973; DePaulo, Rosenthal, Eisenstat, Finkelstein & Rogers 1978; Zaidel & Mehrabian 1969). However, experiments were contrived and artificial by the very nature of separating the information sources. Van de Creek & Watkins (1972) in a series of studies of naturally occurring behaviour concluded that what was said was more important than visual input and that knowledge of demographic information produced as much accurate behavioural postdiction as did exposure to an audiovisual film.

A further difficulty with the channel research was that observers judging the face 'channel' are usually shown, minus sound, facial expressions that occurred embedded in speech (Ekman & Oster 1982). Ekman, Friesen, O'Sullivan & Scherer (1980) found that the relative weight that people gave to facial expression, the tone of speech and body cues depended on the judgement task and the conditions in which the behaviour occurred. There were correlations of judgement made by observers who saw the face without speech and judgements of people who saw the face with speech. But these were quite low on some dimensions (calm-agitated) and high on others (outgoing -withdrawn).

That there is an universality of the facial expressions of emotions has been reliably demonstrated in the later work (Ekman & Oster 1982). Subjects in both literate and pre-literate societies showed the same facial movements when posing particular emotions. But only a small number of facial expressions have been demonstrated to have an universal expression - anger, disgust, happiness, sadness and fear/surprise combined. The differences of expression that occur between cultures are suggested to be due to different display rules that operate within different cultures. American and Japanese subjects sitting alone watching either a stressful or a neutral film showed the same facial movements. But, when an authority figure was present Japanese subjects smiled more and showed more control of facial expression (Ekman 1973; Friesen 1972).

It has also been shown that people can judge static, posed facial expressions (Ekman, Sorenson & Friesen 1969; Izard 1971). They restricted themselves to a small sample of basic emotions and carefully selected photographs of faces that appeared to portray the emotions unambiguously. Taking as a criteria of accuracy the poser's intention, American judges showed almost perfect accuracy. Ekman et al. also obtained judgements from observers in New Guinea and Borneo who had had virtually no contact with Westerners. These judgements were not as accurate as the

American observers but they were above chance and there was agreement between the groups.

The work examining the channels employed video/film techniques, thus was not static, however was still posed or acted. But what is of real interest to a study concerned with nurses' perceptions of affect is spontaneous expressions of emotion that occur in a natural setting. Winkelmayer, Exline, Gottheil & Parades (1978) used filmed interview sequences of 'normal' and schizophrenic people. Both American, British and Mexican observers were able to make judgements about emotion after seeing the films. Perhaps interesting is that the performance of the three groups was equivalent when judging the schizophrenic groups' facial expressions. But the Mexican group were less accurate than the other two when judging 'normals'. This is an isolated finding that has not been replocated (Ekman & Oster 1982). Ekman (1973) in the study of American and Japanese subjects watching a film found similar accuracy between American and Japanese observers. Also, the same people overall were judged correctly by the accurate observers. But it must be unreasonable to extrapolate from this work and suggest that in the real world people from different cultures can readily gauge the emotions of each other on the basis of their facial expressions. Little is really known about the display rules that operate within various cultures. When it is culturally or otherwise desirable to conceal emotion it would be of interest to know if the face evidenced any detectable 'leakage'. Or if a person experiencing emotional upheaval, it is inappropriate to express it, attempts to conceal this by a fleeting expression or by partial expression (Ekman & Oster 1982).

Overall, from the work reviewed so far people have been shown to be able to gauge a limited number of basic emotions such as anger, disgust, happiness, sadness and fear/surprise from facial expression. Much work has used the posed intentional expressions of actors. This small number of basic emotions also appear to be expressed similarly across cultures in terms of facial movement. But much facial expression has little to do with emotion (Ekman & Oster 1982). Little is known about how facial expressions used to punctuate speech (Ekman 1979) or show expressions that are socially required differ from those expressing emotion. There is some evidence for right hemisphere involvement in the spontaneous expression of emotion, but this is at an early stage of research (Ekman, Hagar & Friesen 1981).

There is recent work concerned with identifying attributes of people who can gauge the emotional state of another with accuracy. Also of interest are characteristics of expressive people, individual differences and methods for studying these factors. These areas of work will be reviewed after a discussion of the problem that is central to this area, namely criterion validity. Ekman & Oster (1982) consider this problem in relation to the facial expression of emotion. They ask how it is possible to determine if the information provided by a person's facial expression is accurate? This can be extended to the accuracy of information that could be conveyed by face, voice and body movement. It is necessary to have an independent criterion to verify the presumed link between an emotion and the expression of that emotion.

Methods employed to attempt to achieve criterion validity have not been without problems. Studies that have used stimuli to elicit desired emotional states in subjects such as positive or negative films may well elicit different emotions in different people and possibly a range of negative feelings may be experienced by the same individual watching negative film eg. disgust and embarrassment.

Some studies have employed subject self-report after they have viewed stimulus material. But if the duration of the stimulus exposure was of a reasonable length of time eg. a short film, a range of emotions experienced may not be recalled, not the order in which they occurred. Unfortunately for none of the 'channels', such as voice pitch or body movement, is there evidence that a certain change is indicative of a particular emotional state. Therefore no one channel can serve to reliably validate another. The studies of the neural correlates of facial expression are in their infancy. Ekman & Oster (1982, p.163), suggested that:-

> "Because there is no single, infallible way to determine a persons' 'true' emotional state, it is unfortunate that so few investigations have followed the approach of using multiple convergent measures to gain a more reliable indication of the emotion experienced".

5.4

Inferential Ability

It is now intended to consider work that has been concerned with individual differences in a person's ability to judge another. Most of the work in this area has been in relation to a possible 'trait' of accurate person perception ability, a general ability to judge others. Adams (1927) suggested that good judges of anothers behaviours tended to be egotistic, cold-blooded and more interested in people as tools than as human beings. Much of this work was directed at the ability to judge assumed personality attributes of other people rather than their emotional state. It is described by Cook (1971) and will not be described in detail within this thesis because it bears no direct relevance. Rather, the difficulties encountered in measuring accurate performance will be discussed in some detail as they do have direct relevance to the discussion of the method adopted for this work.

There was considerable controversy about whether a general ability, to judge others in fact existed and if it did, the methods employed to investigate the phenomena had been severely confounded. Cronbach (1955) notes that:-

> "These studies have usually been built around a particular operation in which a Judge (J) "predicts" how another person (O) will respond. Often, for example, both persons describe themselves on a personality inventory, and J is then asked to fill out the inventory as he thinks O did. The extent to which the prediction agrees with O's actual response is taken as a measure of J's accuracy of social perception (or "empathy", "social sensitivity", "diagnostic competence" etc) ".

Cronbach suggested that if the enquiry is interested in identifying correlates of a general ability, such as a personality characteristic, to judge others not only are there two sources of error present in the data but also two types of acçuracy. There could be a constant error in terms of differences in 'level' where a judge typically under/over estimates from the criterion by a fixed amount. There could also be error in terms of 'spread', where judges correctly identify the mean plus the direction of deviation from the mean but over estimates the size of the deviations. The first type of accuracy that he described was 'stereotype' accuracy, an error in correlation. The judge could fail to differentiate and use a mean score ie. identifying an average response. The second type of accuracy was the measure of interest but was obscured by the other three artifacts. However, in an investigation which is concerned with accuracy, itself regardless of type Cook (1971) suggest that:-

> "If we analyse each trait separately - a desirable proceeding in any case - we can use standard scores and so eliminate individual differences in level and spread".

Cline (1964) points out that some judges adopt a strategy of assuming that they are similar to the person being judged. Smith (1957) in a study concerning rapport and social perception in mother/adolescent relationships found that accuracy was due, in part, to the tendency of high rapport mothers to assume similarity with their daughters. This could present further problems in the study of accuracy because when a judge possesses high real similarity with respect to the person being judged, his/her accuracy will correlate perfectly with the degree to which the judge assumes similarity to the person being judged. Accuracy, assumed similarity and real similarity are therefore mathmatically linked. So although some of these early studies were directed toward clinical performance they are difficult to interprete.

Similarly measurement difficulties have been encountered in the studies of therapeutic empathy (Bachrach 1976). In this instance the empathy in question is that of the therapist for client within a psychotherapeutic relationship of a rather more intense relationship than that of nurse/patient in the general ward setting. Empathy being the ability of one person to experientially 'know' what another person is experiencing from their point of view. Truax (1966) had concluded that a therapists empathy could be studied without having heard what the patient said. Raters were able to rate therapists for varying degrees of emphathetic ability given only this information. This was without rater judgements of congruence of therapist response to patient and any outcomes or response measures to treatment. Bachrach (1976) concluded from a multifactorial study of the various components of therpeutic skill that what had been measured by the empathy studies had been a highly correlated complex of attributes associated with the raters' liking for the psychotherapist, and their judgement of 'goodness' of psychotherapy.

A study of nurses' empathetic ability employed the Hogan Empathy Scale (Hogan 1969) to assess empathetic ability and the Barrett-Lennard Relationship Inventory (Barrett-Lennard 1967) to assess patients' perceptions of their nurses' empathic ability (Forsyth 1979). The study found that 50% of the nurses studied were regarded as empathic on the Hogan Scale, whereas 98% of patients in the study saw their nurses as highly empathic. This is regardless of any criticism of external validation ie. therapeutic outcome, and is highly reminiscent of the reported levels of satisfaction with the patient satisfaction studies.

The issues regarding method will be referred to later in the text in relation to the method employed in this study. A very much more recent study tackled the issue of judgement accuracy of health professionals. Gottheil, Exline & Winkelmayer (1979) had noted that in the early literature more distant and dispassionate people were suggested to have been better judges of emotion in another. They suggested that the judgement of the emotional message of a psychiatric patient is a different type of task than judging more explicit behaviour of non-psychiatric subjects. A task that requires different characteristics of the judges. They employed 32 women student nurses training in psychiatric nursing. The students were shown a silent film of 10 'normal' and 10 schizophrenic women each recounting three personal experiences, one of which had made them sad, one happy and one angry. The nurses were tested on a number of measures designed to assess their sociability, specialisation in the physical sciences, emotional adjustment, interests and they had been assessed by nurse supervisors as to their professional skill in relating to patients. A higher score in relating to patients was correlated with successful judgement of the schizophrenic patients, although correlations with total hits did not differ from chance. The personal variables that significantly correlated with more accurate judgement of the schizophrenic women's affect were social involvement, introversion, religious activity, low verbal ability and interest in both the outdoors and music. The characteristics that were related to the successful judgements of the normal women were mechanical interest, social detachment and an interest in sciences.

Overall the 'normal' women were judged more often correctly than the schizophrenic women. However the means were 14.53 and 10.50 respectively and as these appear to have been calculated on the addition of each subjects' mean number of hits for the three emotion conditions this does suggest a low overall hit rate. Although there are problems with validating what the women acutally were feeling when they read the stories, and the correlations described earlier were low, it does appear that judging a 'normal' and a schizophrenic person could be different types of task. However, the degree of schizo-affective behaviour, its type and its duration were not stated by the authors, although they did state that the stimulus pairs had been carefully matched there is a lack of detail. Accuracy of judgement by more socially distant people, was suggested by these authors, to be related to analytic judgement where individuals are judged in relation to previous, accumulated past experiences in terms of stereotypes. Therefore they can be expected to exhibit higher levels of stereotype accuracy. Whereas, where differential accuracy is required more socially involved people do better.

Further individual differences have been reported. Schiffenbauer (1974) found that the emotional state of the observer influenced the emotion they attributed to a facial expression. The degree of expressiveness of the person being judged and observers judgements has been a topic of recent study (Buck, Baron, Goodman and Shapiro (1980). In these experiments those to be observed, 'senders', viewed and discussed a series of emotionally loaded slides (ie. familiar, sexual, scenic, unpleasant and unusual), the observers watched videos of their nonverbal behavour. In order to study the use of cues by observers, the observers segmented the video recordings into meaningful action units ie. they indicated where one piece of information finished and a new one began, denoting breakpoints. Female senders elicited more breakpoints than men overall. The number of breakpoints was strongly correlated with sending accuracy for female senders but not males. The observers saw more breakpoints in the expressions of accurately decodable female senders than poor female senders which was not the case for males. So apparently the more informative women by definition of the number of breakpoints they elicited were more accurately judged. This was not the case for men. When men and women were divided into high and low expressiveness in terms of the degree of accuracy with which they had been judged in the past, the same pattern prevailed in that the more expressive females were judged more accurately than the more expressive males. But the nonexpressive female was judged less accurately than the nonexpressive male. Also, the expressive male sender elicited fewer breakpoints than did the nonexpressive male sender.

This was an interesting carefully conducted study which also indicated a sex difference in the decoding of the judges. Accurate female judges tended to make fewer breakpoints than did less accurate female judges, this was not true of males. Therefore, the suggestion was that there are qualitive differences in the manner in which they receive and convey nonverbal information.

5.5

The Context/Prior Expectations

An interesting American study has been concerned with inferences that nurses make about the degree to which patients suffer (Davitz & Davitz 1981). They make the assumption that a nurse would form her inference of a particular patient's suffering based on what she would expect to feel were she in that situation ie. assuming that the patient would behave in a similar way to herself. The nurses own expectations of suffering for herself were said to be a stereotype based on previous experience of pain and suffering. The context in which suffering occurs being determined in part, by the nurse's expectations, which would also be influenced by diagnosis and other sociodemographic variables.

The study employed a standardised questionnaire measure of nurses' inferences of patient suffering that consisted of a series of vignettes describing 60 different patient profiles. The nurses were required to rate on two scales of 1-7 the degree of physical discomfort or pain and the degree of psychological distress that they would expect the patient described to experience. They also completed a variety of measures of individual characteristics. 94 nurses took part in this aspect of the study, the majority of whom, had been born in the USA. Approximately half had Northern European backgrounds, the remainder represented other backgrounds.

The results demonstrated that nurses who tended to infer relatively high physical pain also tended to report that they had suffered relatively high pain in their own experiences. Inferences of psychological distress were related to ethnic background. Nurses who came from Northern European backgrounds inferred significantly less patient suffering than colleagues from other European or African backgrounds. The suggestion was made on the basis of these findings that:-

> ".... belief systems about the suffering of others are influenced both by an individual's own experiences of suffering and socially learned attitudes about pain and psychological distress. Thus, a nurse who has experienced a great deal of pain associated with a particular injury may believe that another person with a similar injury is also experiencing a high level of pain. But this belief is also influenced by his or her attitudes toward pain, attitudes aquired as a consequence of his or her history of social learning". (Davitz & Davitz 1981).

There were further observations from the correlation matrix that related to individual nurse experience. The number of years of nursing experience was negatively correlated with a measure of stoicism. These nurses reported more stoic attitudes, complained less of psychological distress and tended to prefer less stimulating situations. However, a X^2 analysis had revealed no effect of length of experience on inferences of suffering where the quartile of nurses who had inferred least suffering were compared with the quartile who had inferred most as a function of length of experience.

Davitz & Davitz suggested that those subjects who inferred low levels of patient suffering were employing psychological distance as a defence mechanism to help them to cope with the discomfort that they experience in the face of patient distress. They suggested the method of coping was reinforced by cultural factors such as the accepted norms of emotional expression and stoicism. A second series of studies investigated the possibility that a tendency to cope by use of psychological distance may be reflected by behavioural distance. A behavioural scale was developed and observations were made of nursing behaviour. A study of 26 high inference and 23 low inference Medical-Surgical nurses found that high inference nurses tended to stand closer to their patients, conveyed a warmer emotional tone and touched their patients more. They explained to patients past or future nursing action more often and less frequently dismissed their patient's concerns. However, they did not explore feelings more frequently or offer more sympathetic understanding of the patient's situation. A similar pattern of findings emerged in a paediatric setting but not on an obstetric unit. This would suggest that the situation in which a nurse works has a modifying effect on her behaviour despite predisposition.

However, interesting as this study is in relating beliefs to behaviour, it did not attempt to capture what the nurse knew of how the patient actually felt. It is of interest to note that those nurses who inferred relatively high patient suffering did not do so by exploring their patients feelings - they observed but did not act on this information.

Other work has been concerned with beliefs and expectations nurses have about different classes of patients. In the psychiatric setting nurses have been shown to have a tendency to be socialised into attitude clusters towards diagnostic labels (Anderson 1978; Brady 1976; Wallston, Wallston & de Vellis 1976). A similar effect has shown with cancer patients (Morrow, Craytor, Brown & Fass 1976). These studies suggested that the nurses' beliefs regarding the diagnostic label which may or may not be relevant influenced their predictions of disease progress and outcome.

5.6

Nurses' Accuracy in Judging Patients' Feelings

The literature that is directly related to nurses' perceptions or knowledge of how their patients feel and what they fear is extremely sparse. Carnevali (1966) found that nurses accurately perceived pre-operative patients' fear of the unknown and concern over separation from home. But they did not so readily identify their patients' fears of pain, death and destruction of body image. Carnevali was reporting a series of three small, exploratory studies conducted by different workers. Interview techniques were employed and no statistical analyses were performed. Johnston (1976) asked surgical patients to complete an inventory to describe their worries and simultaneously a nurse completed a similar inventory to indicate how much she knew of the patient's worries. There was poor correspondence between nurse and patient that suggested the nurses had very little information about what concerned patients. The nurses were not only inaccurate, but they showed a strong response bias, identifying almost three times as many patient worries as were reported by patients.

This study has been replicated using a fellow-patient and a nurse as informants (Johnston 1982). The study employed 10 mutually selected pairs of post-operative gynaecological women. The nurses were identified by the patients in terms of the nurse with whom a patient reported most contact. The Hospital Adjustment Inventory (DeWolfe, Barrell & Cummings 1966) containing 22 items that require a yes/no response was again employed. Nurses identified significantly more worries than either patients or fellow patients overall. Fellow-patients correctly identified 49% of patients' worries and nurses 59%, the higher percentage for the nurses reflected the nurses' over-attribution of worries. When the degree of congruence between patients and judges was compared, the fellow-patients were significantly more accurate than the nurses ($p < .02$), although the presence of high levels of false positives were present in both groups of judges. The nurses were most likely to over-estimate worries about discharge, occupation and progress, and to underestimate worries about catching disease, being confined and being liked by staff.

However, as has been mentioned response bias was present for both fellow-patients and nurses. Their false positive rates being 60 and 70% respectively. Where this is uncontrolled the finding of clear superiority for fellow patients must be tentative. The false negative rates were 16% and 30% for fellow-patients and nurses respectively. This is a very interesting study but what remains unknown is the degree of anxiety related to these concerns - there is no reason to assume a simple additive effect. One worry may cause a patient considerably more anxiety than several relatively minor worries. The study employed a yes/no task, therefore items could be marked as yes if they were slightly worrying as well as if they were extremely worrying.

CHAPTER 6

6.1

Clarification of the Research Problem

The literature reviewed suggested that nurses tend not to spend time with patients discussing, even briefly, their fears and feelings. People are able to gauge the basic emotional state of another from their nonverbal behaviour, but this is subject to distortion due to individual differences in expressiveness, display rules, conscious concealment on the part of the person being judged and the complexity of facial expressions which are only partly concerned with expressing emotion. Distortion can also be present within the nurse due to a nurse's prior expectations of a patient's feelings influenced by the nurse's own experience or lack of it, the nurses socio-economic and cultural background, the patient's socio-economic and cultural background and the emotional state of that nurse.

The work of Johnston (1976;1982) suggested that nurses are not aware of what their patients worry about or over-estimate the total number of patient concerns. What has not been explored is when nurses are given the opportunity to estimate the actual degree of anxiety experienced by their patients, rather than an all or none identification of concern, the overall accuracy of that estimate. It is also unclear how well nurses perceive a wider range of feelings such as those concerned with a mood of unhappiness or depression and how this compares with their perceptions of physical feelings such as fatigue or vigor. The emphasis of traditional nurse education has been directed towards physical care and patient well-being. It is not known how well nurses perceive their patients' physical feelings except in terms of the inferences that they made with respect to pain and hypothetical patients (Davitz & Davitz 1981).

It was considered essential at this point to investigate nurses' perceptions of a wider range of potential feelings that patients may experience. Doing this in a natural environment ie. the ward setting, rather than investigating inferences or clinical judgements made about fictional patients. The literature reviewed illustrated well that the measurement of accuracy is a difficult methodological problem. However, the presence of negative affect in hospital patients and its potentially deletereous effect on recovery is an important patient problem. Nurses are in the best position to make clinical judgements about patient affect because of their potential for patient contact. Therefore, attempting to establish the degree of accurate perception of grades of feeling, although difficult, was considered to be important.

6.2

Specific Research Questions

This part of the study sought to establish whether or not:-

1. Nurses were able to estimate the degree of anxiety experienced by patients when they were allowed to express this in terms of the degree of feeling represented by an adjective describing a mood rather than an actual concrete concern eg. financial troubles.

2. Nurses were aware of those patients experiencing unhappiness or depression. Further, were able to estimate the degree of these emotions.

3. Nurses were able to estimate their patients feelings of physical fatigue or physical vigor.

4. Experience in nursing increased the accuracy of nurses' perception of affect.

6.3

Methodological Problems Inherent in Measuring Accuracy of Perception

The methodological problems that were identified in the literature will now be discussed in terms of how they were dealt with in this study.

A. Criterion Validity

There are several problems with using the self-report of emotions or feelings as a valid criterion of the experience of those feelings. It was noted that people may not have a total recall of a range of emotions that they experienced when viewing an evocative film. In terms of this study of patient affect, patients may not have had a total recollection of a range of feelings experienced over the period of a morning and there could have been marked fluctuations over time. However, as no 'channel' such as facial expression, voice pitch, body movement or any other external, observable validation can give a reliable validation of a particular emotion, the patient's self-report was considered to be the closest that could be obtained to an accurate report of that experience.

A rating-scale technique was employed in this study containing a range of possible descriptions that pertained to the feelings being investigated. When the patient was describing his/her experience there were available a number of adjectives that described a similar feeling that could have served as an aid to memory. The provision of a range of similar adjectives could have also helped to minimise differential use of language between individuals because this allowed analysis in terms of their correspondance ie. similarity of use.

The problem of fluctuation of an experience such as an experience of heightened anxiety that was reduced by a critical event eg. a negative test result, was reduced by suggesting a specific time period for the assessment of feelings (eg. lunch time). Thus a criterion that relied upon the use of language and recall from memory over time could be more exact but not perfect.

B. Social Desirability Effect

The effect of a socially desirable response (Edwards 1957) being given was uncontrolled and reduced as much as possible by instructions to subjects andby the size of the patient sample.

C. Scale Artifacts

It was noted in the literature that there were two different types of accuracy present when rating scale methodologies were employed. The first was differential accuracy ie. the ability to differentiate between people and the second, stereotype accuracy, the ability to judge the typical response of a class of respondents. In a study designed to consider the accuracy of nurses' perceptions of patient affect, a coalescence of the two aspects of accuracy to give a global measure was considered to be of interest. A pre-disposition to base judgement on the use of a stereotype was introduced into the design as an additional variable.

The two sources of error that were present with the use of rating scales in terms of response bias ie. level and spread were eliminated by transformation of the scores and use of the standard deviation rather than the mean when the accuracy of subgroups of nurses was being compared.

D. Assumed Similarity

A judgement strategy of assuming that another person would feel and respond in a similar way to oneself was well documented (Smith 1957: Davitz & Davitz 1981). Demographic details of nurses and patients were recorded and the strategy of assuming similarity encorporated into the design as an additional variable.

E. Sex Differences

The possibility of qualitive differences between the sexes in the manner in which they transmit and judge emotional messages was controlled by employing all female judges/nurses who judged an equal number of male and female patients. The effect of the sex of the researcher was held constant. The researcher was always female.

F. Expressiveness

Individual variation in levels of expressiveness was uncontrolled. This was reduced as much as possible by large subject numbers.

Method

The Selection of Measures

The measure of affect selected was a simple self-report mood adjective check list (M.A.C.L) adapted by Lishman (1972) from McNair and Lorr (1964). McNair & Lorr (1964) made the distinction between mood and emotional state, although they viewed them as overlapping categories they suggested that mood states are likely to be more persistent. They regard mood as a concept with the status of an intervening variable. It was conceptualised

> "as an organismic state definable in terms of the antecedent inducing operations and the correlated behaviour consequences".

Intraorganic events such as an illness have been shown to be antecedent conditions inducing mood change as indicated by this MACL (Wilson-Barnett 1977).

The check list had 24 adjectives that refer to five mood factors found by McNair & Lorr (1964). They were tension/anxiety, depression, fatigue, vigour and hostility. Lishman (1972) had reduced the list of adjectives from over 50 to 24 and derived the same mood factors. Wilson-Barnett (1977) in a series of studies involving 355 patients found test-retest reliability and sufficient sensitivity to detect mood changes as a result of an intervention. The original test-retest reliabilities for these factors were of the range r = .61 -.69 (McNair & Lorr). More detail of this measure and its use follows below.

A further measure employed was that of Cognitive Complexity (Bieri et al. 1970). It was noted in the preceding chapter that some people appeared to be more accurate judges of the emotional messages of non/abnormal populations. They appeared to demonstrate superior stereotype accuracy. However, when a schizophrenic population was studied the less socially distant subjects were the better judges. The Cognitive Complexity measure was suggestive of distance in terms of the amount of social information that a person has at their disposal. Bieri et al. (1970, p.162) states:-

> "Cognitive Complexity may be defined as the tendency to construe social behaviour in a multidimensional way, such that a more cognitively complex individual has available a more versatile system for perceiving the behaviour of others than does a less cognitively complex person".

They developed an index based on the extent to which individuals apply constructs differentially when successively categorising a series of acquaintances in a modification of Kelly's (1955) repertory grid methodology. People who sort these acquaintances in a similar way on several dimension are designated cognitively simple, whereas cognitive complexity is attributed to those who sort them differently on each dimension. The cognitively complex infers more social information. Leventhal (1957) found that cognitive complex people predicted less similarity between themselves and others. Adams-Webber (1967) failed to find a significant relation between cognitive complexity and predictive accuracy. However, Adams-Webber (1969) found that relatively cognitively complex people showed more skill than relatively cognitively simple people in inferring the social constructs of others in social situations.

Reiteration of Research Questions

1. Are nurses able to estimate the degree of anxiety experienced by patients?

2. Are nurses aware of those patients experiencing unhappiness or depression?

3. Are nurses able to estimate the extent of their patients feelings of physical fatigue or physical vigor?

4. Does experience in nursing increase the accuracy of nurses' perceptions of affect?

5. Do differing levels of cognitive complexity differentially affect nurses perceptions of affect?

6. When do nurses employ strategies of a) assuming similarity to the patient they are judging and judge as they would feel themselves and b) use a stereotype based on previous experience to make that judgement?

Design

A questionnaire design was employed. The patients completed one questionnaire, the Mood Adjective Check List (MACL) as a measure of their affect or feelings. The nurses completed four questionnaires as measures of:-

1. Their assessment of the patient

2. Their assumption of their own similarity to the patient

3. Their assessment of a stereotypical patient

4. Their degree of Cognitive Complexity

Each nurse assessed two patients, except where this was impossible. All nurses were tested at 2pm having been on duty that morning. They were asked to select two patients who had been allocated to their care that day and assess how they judged that those patients had felt that morning. If they considered there had been a significant change they were asked to use 12 mid-day as a criterion time.

Materials

The Patients' Questionnaire

The MACL comprised 24 adjectives listed on one sheet beside four columns that were headed by a quantitive phase ie. "not at all", "a little", "quite a bit" or "extremely" and were scored 0, 1, 2, & 3 respectively. The sheet was headed by the instruction:-

> "Below are a number of words which describe moods. Please put a cross to indicate how much you have felt the way described today".
> (App. 39).

The Nurses' Questionnaire

1. The MACL identical to that of the patients but with modified instructions ie:-

   > "Below are a number of words which describe different moods. Please put a cross in the appropriate column beside the words to indicate how much you think each word describes how this patient has felt today".
   > (App. 40)

2. The MACL modified with the instruction:-

   > "Imagine yourself in this patient's position. Please rate this list of words as you would expect to feel if you were this patient".
   > (App. 41)

3. The MACL modified with the instruction:-

   > "Imagine a typical patient of the same sex suffering from the same condition. Please rate this list of words as you would expect a typical patient to feel".
   > (App. 42)

These three questionnaires were stapled together in a fixed order where the nurses assessment of the patient was followed by her assessment of herself and finally her stereotype. This was done rather than balancing the order to control for the effect of a self-rating influencing the assessment of the patient and an influence of the stereotype measure on either of the preceding questionnaires.

4. A cognitive Complexity Repretory Grid modified from Bieri et al. (1970). The subjects were presented with 10 roles that would fit different people that they knew. (App.2).

5. The nurses completed a sheet giving brief details of age, marital status, stage of training, ethnic origin and history of hospital admission.

The Pilot Procedures

The nurse questionnaires were tested with a group of six nurses working on the cardiothoracic wards of the hospital employed in the study. Approaches were made to the medical staff by the ward sister and formal ethical committee approval for the study had been obtained. The questionnaires were examined for:-

1. Clarity of instructions
2. Feasibility of completion
3. The face validity of the task

The nurses were seen in pairs at 2pm having been on duty that morning. They were asked to select two patients whom they had nursed for a large part of that day. Two sets of the three questionnaires, one for each patient, were provided and the nurses were asked to record the patient names in the space provided at the top of the first page of each set of questionnaires. The page of personal details was completed and followed by the first and second questionnaire and the cognitive complexity questionnaire. A full explanation was given on completion.

When the nurses had completed the questionnaires the patients were approached. They were asked to complete a short questionnaire as part of a research study and the instructions at the top of the questionnaire were read aloud to them by the researcher.

Both nurses and patients were co-operative and had no difficulty in completing the questionnaire. On inspection the questionnaires were correctly completed. But two points were apparent.

1. It was necessary to answer the questions put by the nurses briefly and as fully as possible but without giving additional information. A full explanation was provided at the end of the procedure.

2. The likelihood of a fluctuation in the mood state of an individual patient during the course of a morning was possible. Therefore, a criterion time of 12 mid-day was selected as being relatively recent in terms of the patients memory of their mood. The nurses would still have contact with their patients prior to preparing hand-over reports and staff lunches.

The Sample

80 student and pupil nurses at various stages of training at one London Teaching Hospital acted as subjects. The age range was 18-35 years. All the nurses were born in the UK and the majority were of British origin. A small percentage (approx 10%) were of West Indian origin.

140 patients were assessed with an equal proportion of:-

a) Male and female
b) Medical and surgical
c) Housed in long open wards and new 8 bedded units

All patients had been in hospital at least one day but had varying degrees of severity of illness and were at varying stages of progress. A small proportion of patients (less than 10%) had an origin which was other than British.

A large London Teaching Hospital was employed in the study. All adult medical and general surgical wards were visited.

The age range of patientswas 15-96 with a mean age of 60 years and a mode of 70-79 years. (App. 43 & 44).

Procedure

Permission was obtained from the ethical committee, the medical and the senior nursing staff to conduct this part of the study. The researcher had visited the ward sisters prior to seeing the nurses in order to identify their availability. Nurses were not aware in advance, of the nature of the research or if and when they would be involved with it. No nurse was seen more than once.

The nurses were all seen by the researcher between 2 and 3pm in a quiet part of the ward or a vacant office. Nurses who had been on duty that morning were involved. They were asked for their voluntary participation and were told that their contribution would be completely anonymous. They then read the following typed written instructions:-

> "Please think of two patients whom you have nursed for a large part of today. Please select patients who are fit enough to co-operate and are not confused. I am now going to give you two sets of three short questionnaires, one for each patient. I would like you to fill them in according to the instructions at the top of each sheet. Complete your patient 01 first and then 02, write the names of your patients in the spaces provided. I shall also be asking the patients to complete a questionnaire. All the completed questionnaires are completely confidential. They will not be seen by anyone other than myself and I do not want you to write your name anywhere on the forms. You all have a subject number.
> Are there any questions?
> If there are any queries at any point please ask me and I will try to answer them.
> Thank you very much for your help".

The nurses then proceded to complete the sheet of personal details which was attached to the top of the set of questionnaires for patient 01. They then assessed their first patient, following this with how they imagined they would feel in the same situation and finally how a typical patient would feel. The same procedure was followed for the second patient. Queries were answered as fully as possible but no extra information for questionnaire completion was given. All the nurses were instructed to use late morning as a criterion time if there had been a change in the patients condition.

When the nurses had completed their assessments the researcher approached the patients and explained that she was a research nurse doing some research with the nurses, and that it would help very much if that patient could complete a short questionnaire. Only two patients were unwilling to co-operate, both were very sick. The researcher excluded another three on the basis of the gravity of their condition.

The researcher read aloud the instructions at the head of the MACL whilst the patient also read them. She then reiterated them, instructing patients to mark the appropriate box beside each word as it best described how they had felt that day. In the case of a change they were instructed to describe how they felt at 12 mid-day. Where patients required assistance the researcher gave that assistance, except in interpreting the words.

Approximately $\frac{1}{3}$ of patients required assistance to record their response. Where this was the case the researcher read the instructions clearly and slowly to the patients followed by each adjective in a standard fashion eg.

> "Does the word shakey apply not at all,
> A little, quite a bit or extremely".

CHAPTER 7

Results

7.1

Introduction to Analysis

The measure used in this study employed an ordinal measure. The previous discussion outlined the considerations of validity that determined the selection of the MACL. However, the problems remained of an ordinal measure of limited range.

Parametric correlational statistics were employed throughout the analysis. Siegel (1956, p26) stated quite categorically:-

> "That parametric statistical tests, which use means and standard deviations (ie. which require the operations of arithmetic on the original scores) ought not to be used with data on an ordinal scale. The properties of an ordinal scale are not isomorphic to the numerical system known as arithmetic".

The concern is with the level of measurement. Only interval and ratio measures possess properties of ascending/decending rank relations between numbers of an equal and known distance between these numbers. The ordinal level of measurement is less high, the numbers alloted to categories have the property of a rank relationship, but the intervals between them on the scale are unequal or unknown.

Anderson (1961) discussed the problem of an appropriate choice of test for a scale employed, and discussed the predisposing conditions for arithmetic operations. He cites

Lord (1953) and contradicts the statement made by Siegel suggesting that:-

> "The statistical test can hardly be cognizant of the empirical meaning of the numbers with which it deals. Consequently, the validity of a statistical inference cannot depend on the type of measuring scale used".

Lord (1953) points out that with a sufficient sample size, numbers tend to behave in the same way regardless.

Non-parametric statistical tests do not make assumptions about the distributions of the samples studied. They do not require normally distributed data sets. Boneau (1960) points out that non-parametric or distribution-free methods have coupled with their freedom from restricting assumptions a disregard for much information contained within the data. They frequently ignore the intercategory differences that do exist between scores. This lack of power can result in a tendency not to reject the null hypothesis when it is actually false.

This problem is exacerbated in the case of a four-point scale and a large sample. There will inevitably be a large number of ties for one or more of the categories and thus much lost information. Boneau (1960) found the t-test to be robust under conditions where there was a violation of the sample distribution assumption. The F-test has also been found to be useful when comparing scores on 5-point scales (Hsu & Feldt 1969).

Visual inspection of the distributions in the current case (App. 45-49) does indicate a skew to the left especially in the patient data. The sample observed reflects the population skew that has been demonstrated extensively elsewhere with different patient groups (Wilson-Barnett, 1977).

Given the large sample size and the problem of extensive ties, 2-tailed parametric tests were applied throughout. The analyses performed were:-

1. Factor Analyses on both patient and nurse MACL data. This was to investigate the similarity of interpretation of the adjectives between the two groups in terms of the derived factors prior to further analysis.

2. An overall comparison of the main factors was made between the two groups. Items within the factors were considered individually rather than as a total score for each factor because of the double correlation that would have been introduced, as well as a loss of sensitivity. Simple and partial correlation coefficients were computed controlling for any effects of cognitive complexity.

3. The patient and nurse data was converted to two transformations (separately for both groups) prior to a breakdown comparison of differential accuracy between nurse sub-groups. This was performed to remove the effect of artificial rating tendencies between individuals, such as an extreme or cautious use of the scale that could result in a spurious correlation (Cook 1971). The rating of male and female patients was compared, first and second year nurses were compared and differing levels of Cognitive Complexity examined in more detail.

   The total data pool was examined on four occasions. Therefore a $p < 0.05$ level of significance was accepted because there was $> 5\%$ likelihood of significance due to chance.

4. The final stage of the analysis involved the MACLs where the nurses put themselves in the position of the patient (Assumed Similarity) and rated a typical patient (stereotype). Partial correlation coefficients were computed between these measures and the nurses rating of the patients. The stage of nurse training and levels of cognitive complexity were examined.

7.2

The Factor Analysis

Factor Analyses were performed on both patient and nurse MACL data. Principal components with iteration and an orthogenal varimax rotation was computed. The factors that were derived were the same as had been found in previous work (Wilson-Barnett 1977, App. 60).

Factor Analysis - Factors Derived from MACL Data on 140 Patients

Factor	Adjectives	Loading	% var	Cum %
1. Vigor	Vigorous Lively Full of pep Active	.75 .705 .85 .73	14	14
2. Depression	Unhappy Miserable Discouraged Depressed	.72 .84 .30 .75	13	27
3. Hostility	Resentful Spiteful Angry Furious	.28 .85 .37 .78	10	37
4. Anxiety	Nervous Tense On edge	.30 .45 .82	8	45
5. Fatigue	Shaky Weary Tired Worn-out	.29 .54 .71 .25	6	51

(App.51 gives a full list of eigen values and percentages)

Factors Derived from MACL Data on 150 Nurse Observations

Factor	Adjectives	Loading	Ratings % Var	Ratings Cum %
1. Fatigue	Shaky	.32		
	Sluggish	.545	15.5	15.5
	Weary	.77		
	Tired	.89		
	Worn-out	.80		
2. Depression	Unhappy	.69		
	Worthless	.33	14	29.5
	Miserable	.84		
	Discouraged	.30		
	Depressed	.84		
3. Vigor	Vigorous	.77		
	Lively	.76	13	42.5
	Full of pep	.80		
	Active	.56		
4. Anxiety	Nervous	.57		
	Tense	.79	10	52.5
	On edge	.76		
5. Hostility	Resentful	.75		
	Spiteful	.31	7	59.5
	Angry	.56		
	(Furious	.11)		
6. Hostility	(Resentful	.15)		
	(Guilty	.17)	6.5	66
	Spiteful	.37		
	Angry	.55		
	Furious	.78		

(App.52 gives a full list of eigen values)

The Factor Analysis from both samples revealed an overall similarity of structure, although there were differences in the rank order. Factor 1 for the patient data and 3 for the nurse data, Vigor suggested a similar understanding of the adjectives concerned as did Factor 4, Anxiety.

Factor 2, Depression, in the case of the nurses, included an additional adjective 'Worthless'. This had a loading of .21 in the patient data and suggested that the word may have had a dissimilar significance for the two groups. The word therefore was not included in an overall Factor structure. The adjective 'Discouraged' had a low loading for both groups (.30 respectively), but did suggest similar use and was therefore retained as part of Factor 2.

The emergence of two Hostility Factors in the nurses data implied that they were not clear Factors and were therefore treated with caution. 'Resentful' and 'Angry' were the most clearly correlated adjectives, although both words had lower loadings in the patients data (.28 & .37). It was decided to retain the factor because of the rank importance evident in the patient data but interpret any findings with caution. The nurse data did suggesting that nurses did not have a concept of a factor of Hostility.

The Factor, Fatigue, included the word 'Sluggish' in the nurse data. In the patient analysis 'Sluggish' had a loading of .14 which indicated that it was not associated with the patient Fatigue factor and was therefore dropped. Shaky appeared to have been understood in a similar fashion and was retained despite lower loadings (.29 & .32).

The final factors that emerged for comparison between the groups were:-

Factor	Adjectives
1. Vigor	Vigorous Lively Full of pep Active
2. Depression	Unhappy Miserable Discouraged Depressed
3. Anxiety	Nervous Tense On edge
4. Fatigue	Shaky Weary Tired Worn-out
5. Hostility	Resentful Spiteful Angry Furious

The remaining adjectives were dropped from the analysis.

7.3

The Main Analysis

The appendices 45 to 49 show the cummulative distributions of the nurses and patients ratings ie. the distributions of the number of patients and nurses rating each word subsumed beneath each factor.

Figure 7.1 shows the total percentage of patients rating each factor as applicable to them (ie. 'a little'/ 'quite a bit'/'extremely') and the percentage of nurses who rated these adjectives as applicable to their patients.

Figure 7.1

Total % Patients and Nurses Rating Adjectives as Applicable

N = Nurses 140, Patients 140

	Patients		Nurses	
Factor 1	Rank		Rank	
Vigorous		21%		23%
Lively		55%		57%
Full of Pep	2	24%	4	29%
Active		51%		69%
		38%		44%
Factor 2				
Unhappy		31%		78%
Miserable		32%		63%
Discouraged	4	21%	3	52%
Depressed		41%		63%
		31%		64%
Factor 3				
Nervous		42%		80%
Tense	3	40%	2	76%
On edge		30%		58%
		37%		72%
Factor 4				
Shaky		51%		69%
Weary		71%		81%
Tired	1	78%	1	87%
Worn-out		40%		67%
		60%		76%
Factor 5				
Resentful		11%		31%
Spiteful		5%		6%
Angry	5	12%	5	33%
Furious		3%		3%
		7%		18%

Figure 7.2

The Ratios of Patients/Nurses Ratings of the Adjectives Relating to Anxiety as a Function of Degree

		Patients		Nurses	Observed Ratings
Nervous	"a little"	1	:	1.7	41 : 69
	"quite a bit"	1	:	2.4	16 : 39
	"extremely"	1	:	3	2 : 6
Tense	"a little"	1	:	1.9	39 : 74
	"quite a bit"	1	:	2	13 : 27
	"extremely"	1	:	1.5	4 : 6
On edge	"a little"	1	:	2.1	28 : 59
	"quite a bit"	1	:	1.6	12 : 20
	"extremely"	1	:	3	1 : 3

Fig. 7.2 shows the ratios of nurses to patients rating each of the adjectives under each level of degree.

Simple zero-order correlation coefficients were computed correlating the patients and nurses ratings of items. Any effect of cognitive complexity was then partialled out. The 5 factors will be presented in sequence preceded by the mean scores.

Factor I Vigor

Table 7.1 Mean Scores Patients/Nurses Ratings

N=140	Vigorous	Lively	Full of pep	Active
Patients	.314	.864	.386	.757
Nurses	.327	.840	.413	1.007

Range 0 - 3

Table 7.2

Simple Correlations of Nurse/Patient Ratings Factor 1

Vigorous	Lively	Full of Pep	Active
.2515 $p < .003$	.145 $p < .088$	.172 $p < .042$	.201 $p < .07$

The correlations between nurses' and patients' adjectives Vigorous, Active and Full of Pep reached acceptable levels of significance. The Partial Correlation coefficient controlling effects of cognitive complexity gave an almost identical result (Table A.1, App. 53).

Factor 2 Depression

Table 7.3 Mean Scores Patient/Nurse Ratings

N=140	Unhappy	Miserable	Discouraged	Depressed
Patients	.479	.521	.293	.593
Nurses	1.087	.880	.673	.920

Range - 0-3

Table 7.4 Correlations of Nurse/Patient Ratings

Unhappy	Miserable	Discouraged	Depressed
.321 $p < .0001$	.281 $p < .001$	.099 $p < .24$	.068 $p < .42$

The correlation between the nurses and patients for the adjectives Unhappy and Miserable reached highly acceptable levels of significance. Discouraged and Depressed failed to approach significance. The partial correlation controlling Cognitive Complexity displayed the same result (Table A.2 App. 53).

Factor 3 Anxiety

Table 7.5 Mean Scores Patient/Nurse Ratings

N=140	Nervous	Tense	On Edge
Patient	.56	.55	.393
Nurse	1.14	1.093	.84

Range 0-3

Table 7.6 Simple Correlations of Nurse/Patient Ratings

Nervous	Tense	On Edge
-.005 $p < .95$	-.0096 $p < .91$	-.038 $p < .65$

None of the correlations of the adjectives subsumed under this factor displayed any agreement between nurses and patients. The same was true when cognitive complexity was partialled out (Table A.3, App 53).

Factor 4 Fatigue

Table 7.7 Mean Scores Patients/Nurse Ratings

N=140	Shaky	Weary	Tired	Worn-out
Patients	.693	.807	1.193	.586
Nurses	.993	1.240	1.353	1.040

Range 0-3

Table 7.8 Simple Correlations of Nurse/Patients Ratings

Shaky	Weary	Tired	Worn-out
.1407 $p < .097$	.1799 $p < .033$	.342 $p < .0001$	.084 $p < .33$

The correlation for the adjective Shaky showed a trend at the 10% level. 'Weary' attained an acceptable level of significance and the correlation for Tired was lightly significant. The correlation for 'Worn-out' was non-significant. Again the same was true for the patient (Table A.4, App. 53).

Factor 5 Hostility

Table 7.9 Mean Scores Patient/Nurse Ratings

N=140	Resentful	Spiteful	Angry	Furious
Patients	.150	.086	.171	.079
Nurses	.390	.087	.287	.053

Range 0-2

Table 7.10 Simple Correlations of Nurse/Patient Ratings

Resentful	Spiteful	Angry	Furious
.172 p .043	.059 $p < .49$	.245 $p < .004$	.189 $p < .025$

The correlations for 'Resentful', 'Angry' and 'Furious' reach the 5% level of significance, 'Spiteful' did not. The partial correlation controlling Cognitive Complexity affected the adjective 'Angry'. The correlation was increased from .2449 to .2774, a small increase of .0325, but this was sufficient to raise the level of significance from .004 to .001. This may be artifactual and will be subject to further scrutiny (Table A.5, App. 52).

7.4

Breakdown Comparison of Accuracy

The Z score variable transformation was performed on both patient and nurse data. Pearson Product Moment Correlation Coeficients were computed for each item subsummed beneath each of the five factors. Only minute discrepancies between the simple correlations on the raw scores and the correlations of the transformed scores were observed (App. 54).

The total patient population was divided into male and female patients. There were 69 women and 71 men, and Pearson Product Moment Correlation Coefficients were computed for the two groups using the transformed data (means for 5 factors, Tables A6-A10, App. 55).

Comparison of Correlations of Nurses Ratings with Male/Female Patients

Table 7.11 Factor I Vigor

	Vigorous	Lively	Full of Pep	Active
Male	.165 $p < .17$	.0165 $p < .89$	.132 $p < .27$	.153 $p < .203$
Female	.372 $p < .002$	.257 $p < .033$	.178 $p < .143$	.234 $p < .053$

The correlations for the adjectives 'Vigorous', 'Lively' and 'Active' reached an acceptable level of significance in the female patient group, although 'Full of Pep' failed to do so. All the correlations of nurses with male patients were non-significant.

Table 7.12 Factor 2 Depression

	Unhappy	Miserable	Discouraged	Depressed
Male	.285 $p < .016$	.353 $p < .003$	.218 $p < .067$	.135 $p < .26$
Female	.365 $p < 002$	.234 $p < .053$	-.032 $p < .79$	.002 $p < .98$

The nurse/patient correlation exceeded the $p < .05$ level of significance with the adjectives 'Unhappy' and 'Miserable' for both male and female patients. In both patients groups 'Depressed' failed to reach significance. The correlation between nurses and male patients for 'Discouraged' did approach the 5% level of significance, this was not the case for female patients.

Table 7.13 Factor 3 Anxiety

	Nervous	Tense	On Edge
Male	.0109 $p < .93$	-.0791 $p < .51$	.0943 $p < .43$
Female	-.0691 $p < .57$	.0540 $p < .66$	-.1406 $p < .25$

None of the correlations for either male or female patients approached significance.

Table 14 Factor 4 Fatigue

	Shaky	Weary	Tired	Worn-out
Male	.2665 $p < .025$	.153 $p < .2$	.375 $p < .001$	.059 $p < .63$
Female	.048 $p < .69$	.214 $p < .08$	.329 $p < .006$	.135 $p < .27$

The correlation between nurses' and patients' ratings for the adjective 'Tired' was significant in both groups. 'Shaky' was significant for male patients. 'Weary' demonstrated a trend towards correlation for the female patients.

Table 7.15 Factor 5 Hostility

	Resentful	Spiteful	Angry	Furious
Male	.227 $p < .057$	−.059 $p < .6$	.2795 $p < .02$	.233 $p < .05$
Female	.084 $p < .5$	.379 $p < .001$	.149 $p < .2$	−.026 $p < .8$

The correlations between nurses' and male patients' scores were significant in the cases of 'Angry' and 'Furious'. 'Resentful' approached significance. For female patients 'Spiteful' was highly significant ($p < .001$).

Nurse/Patient Pearson Correlation Examining the Stage of Training. 1st and 3rd Year Nurses

The stage of nurse training was examined for differential effects of accuracy. 1st year nurses ratings of their patients were compared with those of 3rd year nurses. There were 58 1st year nurses compared with 57 in their 3rd year. (Table A11-A4, App. 56 presents the means for the factors).

Table 7.16 Factor 1 Vigor

	Vigorous	Lively	Full of Pep	Active
1st year	.262 $p < .05$	.3065 $p < .02$	.411 $p < .001$	.339 $p < .009$
3rd year	.169 $p < .2$	.235 $p < .08$	.016 $p < .9$	−.013 $p < .9$

For the 1st year group the correlations between the nurse/patient ratings demonstrated significance in the case of all four adjectives. There were no significant correlations at the 5% level in the 3rd year group, although a trend was demonstrated for the adjective 'Lively'.

Table 7.17 Factor 2 Depression

	Unhappy	Miserable	Discouraged	Depressed
1st year	.307 $p < .02$	.278 $p < .035$	-.029 $p < .8$	.027 $p < .8$
3rd year	.280 $p < .035$	.065 $p < .6$	.129 $p < .34$	.028 $p < .8$

For the 1st year group the correlations between the nurse/patient rating of the adjectives 'Unhappy' and 'Miserable' demonstrated significance, 'Discouraged' and 'Depressed' did not. For the 3rd year group only 'Unhappy' was significant.

Table 7.18 Factor 3 Anxiety

	Nervous	Tense	On Edge
1st year	.073 $p < .6$	-.03 $p < .8$	.042 $p < .75$
3rd year	.047 $p < .73$	.08 $p < .55$	.036 $p < .8$

There were no significant correlations for any of the adjectives subsumed under this factor.

Table 7.19 Factor 4 Fatigue

	Shaky	Weary	Tired	Worn-out
1st year	-.058 $p < .66$	.03 $p < .8$	.348 $p < .007$	-.09 $p < .5$
3rd year	.324 $p < .01$	.264 $p < .05$	.487 $p < .0001$	.119 $p < .4$

The 3rd year group nurse/patient correlations of the adjectives 'Shaky' and 'Weary' were both significant. The correlation of 'Tired' was highly significant, although 'Worn-out' failed to approach significance. The only adjective that was significant in the 1st year nurses was 'Tired' ($p < .007$).

Table 7.20 Factor 5 Hostility

	Resentful	Angry
1st year	.323 $p < .01$	.331 $p < .01$
3rd year	.211 $p < .115$	.174 $p < .2$

The adjectives 'Spiteful' and 'Furious' were omitted from the analysis. No patient in this breakdown of the data had considered themselves to have felt 'Spiteful' and only two patients experienced any degree of 'Fury'. The remaining adjectives 'Resentful' and 'Angry' demonstrated significant correlations between patients and 1st year nurses rating but not between patients and 3rd year nurses.

Pearsons Correlation Coefficients of Nurse/Patient Ratings Comparing Differing Levels of Cognitive Complexity

The total range of Cognitive Complexity scores for all the nurses was 108-259, the possible range was 40-450. Nurse subjects were divided into three groups, the highest scoring $\frac{1}{3}$, the mid-scoring $\frac{1}{3}$ and the lowest scoring $\frac{1}{3}$. The number of subjects in each group and the range of scores was as follows:-

Subjects	Range
36	183 - 259
36	153 - 169
35	108 - 139

The overall distribution of scores can be seen to be skewed to the lower scoring end of the range. The highest scoring group had a within group range of 76 points, the mid-score group a range of 16 points and the lowest 31 points. (Table A15-A19, App. 57, Mean scores for the factors).

Table 7.21 Factor 1 Vigor

	Vigorous	Lively	Full of Pep	Active
High Scores	.006 $p < .97$	-.103 $p < .55$	-.197 $p < .25$	.265 $p < .12$
Mid Scores	.423 $p < .01$	.283 $p < .09$	.531 $p < .001$	.399 $p < .02$
Low Scores	.163 $p < .35$	.039 $p < .8$	.058 $p < .7$	.048 $p < .8$

The only correlations to reach an acceptable level of significance were for the mid-range group of subjects. 'Vigorous', 'Full of Pep' and 'Active' were significant, 'Lively' indicated a trend at the 10% level.

Table 7.22 Factor 2 Depression

	Unhappy	Miserable	Discouraged	Depressed
High Scores	.367 p <.03	.244 p <.15	.105 p <.5	.145 p <.4
Mid Scores	.326 p <.05	.377 p <.02	-.164 p <.3	.044 p <.8
Low Scores	.265 p <.12	.117 p <.5	.120 p <.5	.002 p <.99

The low score group failed to demonstrate significance for any adjective. 'Unhappy' had significant correlation for both high and mid scores and 'Miserable' was significantly correlated in the mid score group.

Table 7.23 Factor 3 Anxiety

	Nervous	Tense	On Edge
High Score	.047 p <.8	-.269 p <.11	-.230 p <.18
Mid Score	.192 p <.26	.175 p <.3	.011 p <.95
Low Score	.079 p <.65	-.035 p <.8	.071 p <.7

No correlation under this factor was of significance.

Table 7.24 Factor 4 Fatigue

	Shaky	Weary	Tired	Worn out
High Score	.332 p <.05	.368 p <.03	.381 p <.02	.148 p <.4
Mid Score	.114 p <.5	.014 p <.9	.363 p <.03	.084 p <.6
Low Score	.188 p <.3	.084 p <.6	.425 p <.01	.123 p <.5

Significant nurse/patient correlations were observed for the adjectives 'Shaky', 'Weary' and 'Tired' for the high scoring subjects. The adjective 'Tired' reached significance among mid and low scoring subjects.

Factor 5 Hostility

The adjectives 'Spiteful' and 'Furious' were again omitted from the analysis because of inadequate variability in sub-groups of patient data.

Table 7.25

	Resentful	Angry
High Scores	.06 $p < .7$	.5025 $p < .002$
Mid Scores	.066 $p < .7$	.454 $p < .005$
Low Scores	.302 $p < .08$	.265 $p < .125$

Significant correlations were obsered for the adjective 'Angry' amongst high scoring subjects and mid scoring subjects. A trend was observed amongst the low scoring subjects and the adjective 'Resentful'.

Table 7.26 is a summary of the correlations that reached an acceptable level of significance for the total patient group and the different sub-groups that were compared.

Summary of Correlations Reaching and Exceeding the 5% Level of Significance for all Groups and Factors

		Male Patients	Female Patients	1st Yr.	3rd Yr.	Hi Scores	Mid Sc	Low Sc	Total Populat.
FACTOR I	Vigorous		.372 $p < .002$	.262 $p < .05$			.423 $p < .01$		.2515 $p < .003$
	Lively		.257 $p < .03$	.3065 $p < .02$					
	Full/Pep			.411 $p < .001$			.531 $p < .001$		.172 $p < .04$
	Active		.234 $p < .05$	.339 $p < .009$			.399 $p < .02$		.201 $p < .02$
FACTOR II	Unhappy	.285 $p < .02$	.365 $p < .002$	.307 $p < .02$	.280 $p < .035$	.367 $p < .03$	.326 $p < .05$		.321 $p < .0001$
	Miserable	.353 $p < .003$	.23 $p < .05$	.278 $p < .04$			.377 $p < .02$		.281 $p < .001$
	Discouraged								
	Depressed								
FAC III	Nervous								
	Tense								
	On Edge								

Table 7.26 (cont)

		Male Patients	Female Patients	1st Yr	3rd Yr	Hi Scores	Mid Sc	Low Sc	Total Popul
FACTOR IV	Shaky	.2665 $p<.025$			.324 $p<.01$	.332 $p<.05$			
	Weary				.264 $p<.05$	.368 $p<.03$			.1799 $p<.03$
	Tired	.375 $p<.001$	.329 $p<.006$	.348 $p<.007$	.487 $p<.000$	.381 $p<.02$	.363 $p<.03$	.425 $p<.01$	.342 $p<.001$
	Worn - out								
FACTOR V	Resentful								.17 $p<.04$
	Spiteful		.379 $p<.001$						
	Angry	.279 $p<.02$				.5025 $p<.002$	.454 $p<.005$		.245 $p<.004$
	Furious	.233 $p<.05$							.189 $p<.03$

For a table of the composition of the subject groups see App. 58. 1st and 2nd year nurses judged an equivalent ratio of female to male patients. There was an equivalent distribution of 1st and 3rd year nurses over the differing levels of cognitive complexity. A tendency was noted for high scoring or less cognitively complex subjects to judge male as opposed to female subjects. The overall ratio of male to female was 55:53, for high scores it was 24:14, for medium scores 15:20 and low scores 16:19. A X^2 failed to demonstrate this as a significant difference between the groups (App. 59). The groups were therefore considered to be discrete.

7.5

Analysis of Judgement Strategies

The analysis of this section was confined to the nurses rating for the patient compared with:-

1. The ratings as if they had been the patient (assumed similarity) and

2. The rating of a typical patient (stereotype)

The untransformed scores were used as the comparison being made was within subjects. Partial Correlation Coefficients were calculated controlling firstly for the effect of the stereotype measure on assumed similarity and secondly the assumed similarity measure on the stereotype.

Table 7.27

Summary Table of Partial Correlation Coefficients of Nurses Ratings for Patients with the Measure of Assumed Similarity Controlling Stereotype and the Stereotype Measure Controlling Assumed Similarity

Adjective	Assumed Similarity	Stereotype
Vigorous	.299 ($p<0.0001$)	
Lively	.419 ($p<0.0001$)	
Full/Pep	.2615 ($p<0.002$)	.3154 ($p<0.0001$)
Active	.426 ($p<0.0001$)	
Unhappy	.185 ($p<0.03$)	.2785 ($p<0.001$)
Miserable	.293 ($p<0.0001$)	.222 ($p<0.009$)
Discouraged	.293 ($p<0.0001$)	.254 ($p<0.003$)
Depressed	.277 ($p<0.0001$)	.246 ($p<0.003$)
Nervous	.232 ($p<0.006$)	.248 ($p<0.003$)
Tense		.398 ($p<0.0001$)
On edge		.3045 ($p<0.0001$)
Shaky	.328 ($p<0.0001$)	.240 ($p<0.004$)
Weary	.423 ($p<0.0001$)	.2195 ($p<0.009$)
Tired	.391 ($p<0.0001$)	.389 ($p<0.0001$)
Worn-out	.2725 ($p<0.001$)	.392 ($p<0.0001$)
Resentful		.243 ($p<0.004$)
Angry		.395 ($p<0.0001$)

N = 140

Table 7.27 gives the partial correlation coefficients between the nurses' ratings of the patients and their ratings of Assumed Similarity and the stereotype. Only those correlation coefficients reaching the 5% level of statistical significance are reported (for a complete table see App. 59).

The nurses' rating for the patient of the adjectives relating to Physical Vigor, demonstrated correlation with the measure of Assumed Similarity with the exception of descriptor 'Full of Pep'. The ratings of this descriptor correlated with both the other ratings.

The reverse was present with the rating for the patient of the adjectives relating to Anxiety and Hostility. The rating for the patient of these adjectives correlated with the ratings of the stereotype measure. This was with the exception of one adjective 'nervous' which correlated with both the other ratings.

The rating for the patient of the adjectives relating to Depression and Fatigue correlated with both the Assumed Similarity and the stereotype measure.

Comparison of Groups of Nurses and Their Ratings of Patients, the Assumed Similarity Measure and the Stereotype Measure

Nurses were again divided into groups, firstly according to their stage of training and secondly their level of cognitive complexity.

Stage of Training

Table 7.28 Summary of Partial Correlation Coefficients of Nurses Ratings for Patients with the Measure of Assumed Similarity Controlling Stereotype and the Stereotype Measure Controlling Assumed Similarity by Stage of Training

1st year N = 58, 3rd year N = 57

Adjectives	Assumed Similarity		Stereotype	
	1st year	3rd year	1st year	3rd year
Vigorous	.348 ($p < 0.008$)			
Lively	.353 ($p < 0.007$)	.445 ($p < 0.001$)		
Full/Pep	.439 ($p < 0.001$)	.32 ($p < 0.02$)		.630 ($p < 0.0001$)
Active	.439 ($p < 0.001$)	.413 ($p < 0.002$)		
Unhappy	.414 ($p < 0.001$)			.428 ($p < 0.001$)
Miserbale	.44 ($p < 0.001$)	.377 ($p < 0.004$)		
Discouraged	.456 ($p < 0.0001$)			
Depressed	.307 ($p < 0.02$)			.289 ($p < 0.03$)
Nervous	.459 ($p < 0.0001$)		.273 ($p < 0.04$)	.310 ($p < 0.02$)
Tense			.372 ($p < 0.004$)	
On Edge			.2905 ($p < 0.03$)	
Shaky		.367 ($p < 0.005$)		
Weary	.546 ($p < 0.0001$)			.415 ($p < 0.001$)
Tired	.469 ($p < 0.0001$)	.293 ($p < 0.03$)		
Worn-out	.348 ($p < 0.008$)	.396 ($p < 0.002$)		
Resentful			.336 ($p < 0.01$)	.348 ($p < 0.009$)
Angry				.2965 ($p < 0.03$)

Nurses in their 1st and 3rd year of training were separately analysed. Table 7.28 gives the partial correlation coefficients between 1st and 3rd year student nurses' ratings of their patients and their ratings of Assumed Similarity and the Stereotype. Only those correlation coefficients reaching the 5% level of statistical significance are reported (for a complete table see App. 60).

The overall picture that emerged was that 1st year nurses demonstrated a more consistent correlation between their ratings of patients and their ratings of the Assumed Similarity measure. Of the 17 adjectives that were rated for the patient, 11 correlated with the rating of Assumed Similarity compared with 4 with the stereotype measure. The 3rd year group presented a more complex picture where there were an equal number of significant observations with both measures (7 & 7).

Levels of Cognitive Complexity

Partial correlation coefficients were again computed between the nurses' ratings for the patients and the two measures.

The highest scoring subjects (N=36) with a range of 183-259 were compared with the lowest scoring subjects (N=35) with a range of 108-139.

Table 7.29 Summary Table of Partial Correlation coefficiants of Nurses ratings for Patients with the Measure of Assumed Similarity Controlling Stereotype and the Steroeotype Measure Controlling Assumed Similarity by High and Low Cognitive Complexity Scores
(for complete Table see App. 61)

High scores N=36
Low scores N=35

	Assumed Similarity		Stereotype	
	High Scores	Low Scores	High Scores	Low Scores
Vigorous		.508 ($p<0.002$)		
Lively	.495 ($p<0.003$)	.5395 ($p<0.001$)		
Full/Pep	.4775 ($p<0.004$)			.743 ($p<0.0001$)
Active	.4455 ($p<0.007$)			.354 ($p<0.04$)
Unhappy	.387 ($p<0.02$)			.3845 ($p<0.025$)
Miserable		.374 ($p<0.03$)		
Discouraged				.359 ($p<0.04$)
Depressed	.517 ($p<0.001$)			.369 ($p<0.03$)
Nervous	.402 ($p<0.02$)			.481 ($p<0.004$)
Tense				.367 ($p<0.03$)
On Edge			.375 ($p<0.03$)	.380 ($p<0.03$)
Shaky	.324 ($p<0.05$)	.367 ($p<0.03$)		
Weary		.499 ($p<0.003$)	.564 ($p<0.0001$)	
Tired	.599 ($p<0.0001$)	.422 ($p<0.01$)		
Worn-out	.456 ($p<0.006$)			.595 ($p<0.0001$)
Resentful				
Angry			.526 ($p<0.001$)	.437 ($p<0.01$)

The overall pattern of correlations suggested a tendency for subjects with high cognitive complexity scores to rate patients as they rated themselves (Assumed Similarity) more often than rating them as they perceived a typical patient (stereotype) would feel (9 & 3). Whereas, there was a tendency in the opposite direction for the lower scoring subjects (6 & 10).

CHAPTER 8

Discussion of Results

8.1

The Factor Analysis

The factor analysis was performed to identify the mood factors contained within the MACL and to establish a consensual understanding of the language used to express them by both the patients and the nurses.

The factors derived from the patient data were identical to those derived by Wilson-Barnett (1977). But there was dissimilarity in terms of the number of adjectives subsumed under one factor, Depression. Wilson-Barnett had identified eight adjectives related to and subsumed under Depression (App. 50). Whereas, the present study identified only four in the patient data and five in the nurses' data. The reason for this would be a result of comparing two heterogeneous sets of patient data. The words found to be unrelated in this set of patient data were helpless, guilty, hopeless and worthless.

The component emotions of self-disgust and contempt, and guilt were found by Izard (1972) to characterise pathological manifestations of depression. Zung (1973) described depression as a human experience extending from normal mood swings to a pathological state. The factor analysis of this patient population suggested that the self-deprocating feelings experienced by many pathologically depressed patients were not being experienced. To take this a stage further and suggest that the current population did not contain any patients whose feelings of depression were similar to those described as pathological would be unjustified. But, the factor analysis suggested that psychopathological depressive states were not strongly represented. It is of further interest that the adjectives relating to feelings of self-deprocation and guilt were not related to the nurse factor, Depression. This would suggest, either that nurses without specialist knowledge are unaware of the association between depression and feelings of self-deprocation, or that the nurses accurately perceived that none of their patients were experiencing this depth of depression.

The lack of one clear factor representing Hostility within the nurse data could also have been indicative of unfamiliarity on the part of the nurses. Feelings of resentment and anger are not associated with the hospital experience in the teaching during nurse training. The nurses in this study were all student nurses and, therefore, had had less than three years experience in nursing. The experience and maturity in the job necessary to objectively appreciate that patients can feel resentment and anger about their illness and situation may take rather longer to develop. But, this interpretation is complicated by the small number of patients who reported feelings of hostility in this study (App. 49). Overall, less than 10% of patients expressed such feelings. Despite these observations this factor was retained for the analysis because it was considered to be of interest to observe whether nurses could identify those patients who had reported Hostility.

With the exception of the factor Hostility which has been qualified, the adjectives that were retained for the analysis had shown a clear association with the four factors which they identified by both the patients and nurses. This implied an equivalent understanding and use of language by both groups. Those words that showed no common understanding were not included in the analysis.

8.2

The Main Analysis: Overall Observations and Global Sensitivity

The overall profile of feelings of anxiety and depression were similar to previous work (Wilson-Barnett 1977). The actual proportion of patients reporting feelings relating to depression was larger in this study than had been reported in the studies of Maguire et al. (1974) and Moffic et al. (1975). But these two studies reported the incidence of pathological depression that they believed to be present, and had employed a psychiatric assessment interview in the method and were, therefore, not directly comparable with this study. The present study was concerned soley with the expression of feelings and when the four adjectives were considered together this represented slightly less than $\frac{1}{3}$ of patients (31%).

Slightly more than $\frac{1}{3}$ of patients (37%) reported feelings of anxiety and the nurses reported anxiety for 71.6% of their patients and feelings of depression for 64%. This illustrated a consistent tendency on the part of the nurse to over-estimate the negative effect experienced by their patients by 100%. Johnston (1982) observed that nurses identified two thirds more patient concerns overall than had been identified by those patients. The degree and intensity of anxiety implied by these worries had not been investigated within her study. Figure 7.2 shows the ratio of nurse to patient ratings of each of the adjectives according to their degree of intensity. The adjective nervous shows an ascending over-estimation by nurses as the intensity denoted by the scale increases ie. 1:1.7, 1:2 and 1:3. A similar profile emerged for the adjective On Edge, but there was less variation of rating for the adjective

Tense. But, when inspecting the small numbers of patients rated by nurses as 'extremely' anxious in terms of any three of the adjectives, and effect of a systematic tendency to differentially over-estimate was very slight (App. 45). These observations concur with those of Johnston that nurses tend to over-estimate in a general fashion as to the total amount of anxiety experienced by hospital patients. The profile for the rating of the adjectives denoting depression was similar (App. 44).

By contrast to the large, global over-estimation of patient negative affect by the nurses, the ratings by nurses of their patients' physical feelings were over-estimated to a much lesser extent. The amount of physical fatigue and vigo. reported as being present by this population of patients was more similarly reported by the nurses on behalf of the population (App. 43 and 46).

Physical feelings of fatiguewere the feelings reported by a higher proportion of patients overall. They were the highest in rank importance for both the patients and nurses (Fig. 7.1). The second most important in rank order for the patients was vigour followed by the factors denoting emotional affect. The rank order of the nurse factors mirrored that of the patients factors except for the factor vigour, implying not a simple under-estimation by the nurses of the overall level of vigour experienced by their patients, but rather a gross over-estimation of anxiety and depression.

8.3

The Main Analysis: Correlation Coefficients Between Patients' Ratings and Nurses' Ratings for those Patients

Individual adjectives have been correlated rather than combining the adjectives under each mood factor. For this reason the discussion will focus on 'feelings' as opposed to moods. That the individual adjectives represent the moods has been demonstrated by the factor analysis. But the discussion of feelings as opposed to the more absolute state of a mood allows the discussion of situations where the correlational pattern is incomplete. Within the analysis of this section of the study a mood is viewed as a construct and those adjectives or feelings that are associated with the mood are viewed as dimensions.

In general the simple correlation coefficients observed were of a low order ranging from $r = -.005$ to $r = .342$. The range of correlations that achieved an acceptable degree of statistical significance at the 5% level was $r = .172$ ($p < 0.042$) to $r = .342$ ($p < 0.0001$). In order to understand fully the low order correlations observed in relation to interpreting their meaning, it is necessary to consider the nature of the constructs that are represented by each of the five factors. Taking as an example the construct Depression, the factor analysis identified that the words unhappy, miserable, discouraged and depressed were used in association with each other in a relatively consistent fashion. These words were held to imply the construct depression on an individual basis for the purposes of the correlational analysis ie. each pair of patient and nurse adjectives were correlated individually. There was no one-to-one relationship between that construct depression and a particular adjective. For example, to signify the construct depression a patient may emphasise the word unhappy and, to a lesser extent, the word miserable, whereas the nurse may reverse this. Considering the patient to be depressed she may emphasise the words miserable and depressed and, to a lesser extent, unhappy. Although both parties have recognised the presence of feelings of depression, correlations of a relatively low order would result because of this differential and uncontrolled use of language. The large sample of 140 patients would allow low order correlations to be statistically significant. It was decided, that because of the difficulty of the task with 1 in 4 chance of exact correspondence between adjectives used to describe the construct, to consider correlation coefficients of the order of $r = .2$ or more as meaningful for further interpretation.

The first specific research questions posed during this study were concerned with whether or not nurses were aware of those patients experiencing anxiety or depression, and whether or not they could estimate the degree of those feelings of negative affect. In the case of anxiety there were no correlations between the patients and nurses ratings of any of these adjectives. It has already been discussed that the nurses tended to over-estimate patient anxiety in a global and non-discriminating fashion this was reinforced by the absence of correlation.

A global picture of a general and undiscriminating over-estimation of patient depression by the nurses had also been observed. This was confirmed by the lack of correlation present between nurses and patients ratings of the adjectives discouraged and depressed. But, the correlations between the nurses and patient ratings of the adjectives unhappy and miserable were $r = .321$ and $r = .281$ respectively ($p < 0.0001$ and $p < 0.001$). Inspection of the cumulative distribution suggests that these observed correlations were, in fact, artifactual (App. 44). The nurses had rated 77.8% of patients as unhappy and 62.85% of patients as miserable. The correlations had been achieved by rating a majority of patients as 'a little bit' unhappy or miserable and, therefore, included those patients who had so rated themselves.

The third question posed was whether or not nurses were able to estimate the extent of their patients' physical vigour or fatigue. Inspection of the correlations between patients' and nurses' ratings of the adjectives subsumed under the factor Vigor indicated that nurses were aware of how vigorous and how active their patients felt (Table 7.2). The adjectives 'vigorous and 'active' demonstrated correlations of $r = .2525$ and $r = .201$ respectively and inspection of the frequency distributions suggests that these were meaningful correlations (App.43). The ratings for the adjective lively failed to correlate and the correlation of the ratings of the descriptor 'full of pep' was of an unacceptably low order. The nurses did have some idea, as demonstrated by vigorous and active as to the degree of physical vigour their patients had felt.

Table 7.8 shows the correlation coefficients between the patients and nurses ratings for the factor Fatigue. One of the adjectives, tired, demonstrated a highly significant correlation of agreement between the nurses and patients ratings ($r = .342$, $p < 0.0001$). That the nurses had a clear idea of how tired their patients felt was further indicated by the frequency distribution of scores (App. 46).

Turning to the final factor Hostility, the only adjective to reach the criteria of acceptability was angry ($r = .245$, $p < 0.004$) Inspection of the frequency distribution suggests that this was an artifactual observation. The nurses corresponded with those patients who had rated themselves as angry by over-estimating a sufficient proportion of patients to include those patients (App. 47). The increase in the size of the correlation as a result of partialling out the effects of cognitive complexity must therefore be meaningless. Cognitive complexity was not found to be influential in this undifferentiated form and will be discussed in relation to the analysis that deals with specific comparisons of nurse sub-groups.

The results did indicate that nurses knew which of their patients were feeling relatively vigorous and how active patients felt that they had been. They also had a clear idea of who felt tired and the degree of this tiredness. They were, however, unaware of the actual extent of feelings of anxiety amongst their patients and the correlation coefficients indicated that they were also unaware of who was feeling anxious. The picture was the same for feelings of depression.

The nurses' inaccuracy of judgement of the negative affect experienced by patients can be accounted for by several factors. There are qualitative differences in the potential for expression between feelings of a physical nature and those of psychological nature. Physical tiredness is spontaneously expressed by most people. It can also be readily inferred by the amount of activity that a patient feels able to tolerate and the number of rests that a person is observed to take. Feelings of physical vigour are also observable in terms of a relative increase in activity and the expression by the patient of feeling better.

Feelings of Anxiety and Depression are more complex. A depressed person may complain of feelings of tiredness as a result of that depression, but may be reluctant to express feelings of unhapiness or misery unprompted. The feelings associated with depression can be less tangible and may well not be recognised by the patient for what they are. Anxiety, similarly, may be something patients are less willing to express unless prompted or encouraged. But, this aspect of the discussion must be speculative as little is known about how much hospital patients feel able to express about their negative feelings. What is known is that there are cultural differences regarding the expression of feelings and Northern Europeans express fewer than do other cultures. Further, specific behaviours associated with a negative affect such as anxiety are more difficult to interpret than those associated with purely physical feelings. The anxious patient expressing that anxiety through complaint may be viewed by the nurse as simply difficult. Further, the nurses' behaviour in actively ignoring those who complain can save to discourage the expression of that affect (Stockwell 1975).

The ways in which nurses could gain knowledge about their patients' mood or feelings was discussed in the literature review for this part of the study (Chapter 5). The area of the accuracy of non-verbal communication proved to be complex, but the human face gave the most information in terms of accurate recognition of very simple emotions eg. anger, disgust, happiness, sadness and fear/ surprise. But it was noted that much facial expression has little to do with emotion (Ekman & Oster 1982). Ekman et al. (1980) found that the relative weight that observers gave to facial expression, the tone of speech and body cues depended on the judgement task and the condition in which the behaviour occurred. In this context the nurses' ability to make use of these cues would depend, firstly on her knowledge of their significance and secondly, her expectations of their occurrence.

Nurses are not taught routinely about non-verbal behaviour associated with particular emotional or mood states. Nor is there any formal teaching, with the exception of one or two nurse education establishments, about the emotional responses of patients to illness and hospitalisation. Further, it was noted in chapter 5 that the patterns of verbal communication between nurses and patients do not allow much opportunity for nurses to gain this information. The studies of cancer and gynaecological patients suggested that nurses tended to avoid discussion of feelings (Pepper, 1977: Bond, 1978). Further, Macloed Clark (1983) found that nurse enquiries as to how patients felt occurred rarely, and when they did occur were posed in the form of a closed question designed to elicit a desired response. It is difficult to generalise from three relatively small and specific studies to the total nurse population in the U.K. But, these studies indicated a relatively low emphasis on communication per. se. on the part of nurses and a higher emphasis on the nursing tasks in need of completion. This reflected a realistic appreciation on the part of these nurses of the importance of physical care but a lack of appreciation of the importance of communication.

There is a strong tradition of nurse education that has emphasised the physical consequences of illness and disease. It is not intended to in any way suggest that the teaching about the physical consequences of illness is unimportant, rather that the results of this part of the study do reflect the lack of teaching of psychological consequences and communication skills. The nurses' ability to estimate the degree of 'tiredness' experienced by patients and their feelings of physical vigour reflect, in part, the emphasis of nurse education. But, the differences between judging physical and psychological feelings complicate this as a simple conclusion.

The artifactual correlations observed between the nurses' and patients' ratings of the adjectives unhappy and miserable do suggest the novelty for this group of nurses of thinking about feelings of depression in their patients. It was apparent that the strategy adopted by the nurses was to rate the majority of patients as feeling 'a little bit' miserable and 'a little bit' unhappy. This suggested that these nurses had not considered feelings of depression to be relevant for most of their patients. It was the demands of the rating task imposed by this study that elicited this response.

The over-estimation of the overall presence of patient anxiety was a replication of previous findings (Johnston, 1982). The nurses in Johnston's study were shown to have been aware of some of their patients' worries. The over-estimation coupled with the lack of correlation observed in this study suggest that the nurses expected a majority of patients to feel some anxiety, but were unaware of who was affected and to what degree. These observations will be discussed further in relation to the judgement strategies of assuming similarity to the patient being judged or adopting a stereotype.

Having considered the opportunities that nurses have for assessing the negative feelings of patients and the educational preparation they receive for that task, it is now intended to turn to their ability and strategies for coping with the negative feelings of patients. Davitz and Davitz (1981) found that when a fictional patient vignette was being judged, Northern European nurses tended to infer less patient suffering than did their Southern counterparts. Further, when two groups of high and low inference nurses were compared, the low inference nurses tended to be more physically distant when nursing patients. This was interpreted as a coping strategy of distancing and avoidance.

The Davitz & Davitz study lacked a study condition that considered the accuracy of the nurses' inferences of actual patient suffering.

This part of the current study was primarily concerned with the nurses' judgemental accuracy of their patients' suffering expressed via the negative feelings of anxiety and depression. Although accuracy was not observed, the observation of over-estimation of patient Anxiety is a direct contradiction to the Davitz & Davitz findings. The low inference nurses were assigned to that category as a result of comparison with the other nurses. The likelihood exists that had the nurses' inferences of suffering been compared with patients' reports of suffering, both groups of nurses would have over-inferred the extent of that suffering. The suggestion of a coping strategy of distancing and avoidance appears less obvious in the present study where the nurses appeared able to infer the presence of Anxiety and Depression for a majority of their patients. But, the relatively undifferentiated ratings of these feelings in their patients still suggests psychological distance. An interesting twist in this study is that the nurses who were inferring overly high levels of anxiety were judging predominantly Northern European patients who also may be pre-disposed to under-infer suffering. The problems of self-report of feelings will be discussed in a later section.

Turning now from psychological feelings to physical feelings, the nurses' ability to gauge the physical well-being of patients in terms of their tiredness and their feelings of vigour that was observed in this study has important implications for nursing care. A large proportion of care that a nurse gives to her patients is physical care and recovery is associated with the attainment of physical milestones. The success of a nurse mobilising a patient is dependent upon knowledge of that persons physical limitations. A nurse who knows when a patient is feeling relatively vigorous or is feeling tired can intelligently plan mobilisation in order that it is within the capacity of that patient. That tiredness is a more straightforward feeling to gauge in a patient, in terms of expectation of tiredness as a result of disability, does not make this any the less useful a finding. The need for an adequate provision for rest and sleep is important in the process of recovery and rehabilitation (Wilson-Barnett & Fordham 1982).

8.4

Breakdown Comparison of Accuracy Between Patient and Nurse Sub-Groups

Male and Female Patients

The correlation of the adjectives between nurses ratings and those of male and female patients for the factors Depression and Fatigue failed to show any different patterns. However, both Vigor and Hostility did demonstrate a degree of differential accuracy. For the factor Vigor, female patients were perceived more accurately on three of the adjectives, vigorous, lively and active (Table 7.26). Female patients expressed lower levels of vigour overall than did male patients and the nurses appeared to recognise this (App. 55). The implication is that the nurses under-estimated the degree of Vigor that the male patients reported feeling.

An interesting differential accuracy was observed for the factor Hostility. Male patients were more accurately perceived as feeling angry or furious and female patients were perceived as feeling spiteful when they reported that feeling. These observations are directly supported by the literature about sex role stereotype, where men are permitted to feel and express anger and spite is perceived as a female feeling (Rosenkrantz, Vogel, Bee, Broverman & Broverman, 1968). This was the only clear sex-appropriate expression of feeling that was differentially perceived by the nurses.

The observation of differential accuracy in terms of female levels of expressed Vigor will be discussed further in relation to the use of judgement strategies.

First and Third Year Student Nurses

The research question posed was concerned with whether or not the accuracy with which nurses perceived how their patients feel increases with experience. However, the first year nurses had a clearer notion than their third year colleagues of their patients feelings of Vigor (Table 7.26). Each adjective subsumed under this factor achieved a statistically significant correlation between the first year nurse and patient ratings. The third year nurses did not show any correlations between their ratings and the patients' ratings of Vigor. But, the third year nurses did demonstrate a higher degree of perception of patient fatigue. The ratings between these nurses and their patients correlated for three of the four adjectives. This was a stronger observation than had been the case for the total group. The observations relating to Factor 3, Depression, can again be seen to be due to inclusion of a large number of patients as a 'little bit' unhappy or miserable (App. 62 & 63).

These observations suggested that in the initial stages of nursing experience nurses are more aware of patients' feelings of well-being. But as experience increases this awareness diminishes. The reason for this could well be due to a gradual sensitisation of these nurses to the physical consequences of disease that results in an inability to recognise well-being when it is present. A problem that may occur where there is a failure to recognise well-being could be in terms of recognising progress during recovery. But, the third year nurses had an accurate overall estimate of Fatigue amongst patients and this could imply that they would recognise diminishing fatigue when that occurs. The necessity for this knowledge has been discussed in relation to rehabilitation.

Different Levels of Cognitive Complexity

The nurse and patient ratings were compared for three groups of nurses having low, mid-range and high cognitive complexity scores. These imply relative cognitive complexity, medium complexity and relative cognitive simplicity respectively. The two groups of relatively less complex nurses achieved more correlation between their ratings and their patients ratings than did the relatively complex group (Table 7.26). When the factor Depression was excluded the medium and less complex presented a similar picture with the less complex demonstrating perception of their patients' feelings of Fatigue and the medium group perceiving feelings of Vigor.

What was of more interest was that these nurses who were relatively more complex achieved one single correlation between their ratings and those of their patients. This was with the adjective tired, which has been shown to be the most robust overall. The slight increased accuracy of judgement shown by the medium and less complex nurses can be explained in terms of the more complex. The complex nurses were more inclined to differentiate between patients, any general responses of people in hospital could be lost. Taking the factor Fatigue, the less complex nurses were able to detect consistency between the use of the adjectives that served as dimensions of this factor. This was not the case for either the medium or more complex nurses. In general the over-differentiation of the more complex nurses lead to a loss of accuracy.

The socially involved nurses of Gottheil et al. (1978) were more perceptive in gauging the feelings of schizophrenic women from their less predicable non-verbal behaviour. However, those nurses who were less socially involved were more perceptive in gauging the feelings of 'normal' women. This suggests a degree of equivalence where detachment is equated with the use of fewer constructs ie. the more complex nurses behaved in a socially involved fashion and inferred more constructs than were present.

8.5

The Use of Judgement Strategies

In general the observations derived from the measures where the nurses put themselves in the patients' position Assuming Similarity to that patient and, imagined a Stereotype Patient were far from clear. For the factors Depression and Fatigue the Assumed Similarity measure and the Stereotype measure were both correlated with the nurses ratings of their patients.

The other three factors were treated differently. For the factor Vigor, the nurses' ratings for the patients correlated predominantly with their ratings of themselves as if they were that patients. This applied to three of the adjectives subsumed under this factor but not to the descriptor 'full of pep'. The lack of a clear rating pattern with this descriptor could well indicate the apparent lack of likelihood of nurses considering this to be a feeling that many people would experience as patients.

The observation of the nurses' Assuming Similarity to the patients when rating overall levels of vigouris of interest in relation to their overall increased accuracy when judging the degree of vigour reported by female patients. The group of nurses acting as subjects were all female and were able to gauge the feelings of vigour of female patients more clearly than males. As this strategy was adopted by the majority of nurse subjects, there was unlikely to have been a link between Assumed and actual Similarity present in the data. The female patients were from a widely differing range of ages and social background. The suggestion is, therefore, that in terms of feelings of well-being, when a hospital patient, females are likely to respond in a similar fashion overall, and this can be judged by a proportion of female nurses. this interpretation must be qualified, however, by the third year nurses inability to gauge feelings of well-being.

The factor Hostility presented a picture of the use of a stereotype by the nurses, as if they knew that they would not feel like that, but considered that a typical patient may feel that way. The adjectives 'Spiteful' and 'Furious' had been omitted from further analysis so no further comments can be made about them. But the adjective 'Angry' had been found to be correlated when nurses and male patients ratings were compared. This further reinforces the notion of the use of a stereotype that was sex typed.

The factor Anxiety also presented a picture of the use of a stereotype by the nurses with the exception of the adjective Nervous. This again suggested different expectations for the nurses of themselves than of patients. Given the observed generalised over-estimation of Anxiety by the nurses, this further reinforces their expectation for hospital patients to feel very anxious. The nurses had only rated the adjective nervous as they would have expected to feel.

It is not possible to interpret the observation for the factors Depression and Fatigue due to methodological problems present in the design of this part of the study. This will be expanded in the section discussing the method problems in this part of the study.

First Year Nurses Compared with Third Year Nurses

An overall pattern emerged where first year nurses presented a clearer picture of rating themselves in the patient's position as they had rated the patient, rather than rating the stereotypical patient as they had rated the patient with an observed ratio of 1:3. They were three times more likely to consider that their patient had felt as they imagined they would have felt. The third year nurses presented a picture of being equally likely to rate the patient similarly to themselves or to a stereotypical patient (7:7). The general picture was one of progressing awareness of differences between people and the development of expectations based on experience, but the differences between the two groups was not subjected to statistical testing and will be discussed no further. The observation of a stronger tendency for both the first and third year nurses' ratings of patients to correlate with their ratings of themselves is perplexing given the overall correlational profile. But this again by a reflection of distortion due to problems with the method employed.

Levels of Cognitive Complexity

The summary table of the overall correlation pattern of the nurses ratings of their patients and the two measure of judgement strategy does suggest a difference in strategy use between the relatively complex and those who were relatively less complex. Those nurses with high scores, the relatively less complex were more inclined to rate the patient in a similar way to themselves at the ratio of 10:3. The low scores or more complex subjects did not present such a decisive picture but they more inclined to rate the patient as they had rated a stereotypical at the ratio of 10:6. The pattern of the correlation present does suggest that there could be an effect present that is worthy of further study. It is likely that those people who are relatively more complex are less inclined to adopt a judgement strategy of assuming others to respond as they would respond. But no assumptions of the accuracy of a stereotype held could be assumed, as the constructs of that stereotype could be over-differentiated

8.6

Methodological Problems

The first major short-coming in a study of this type is in the crudeness of the measures employed. The M.A.C.L. was selected because of validity inferred from the original development of the measure by McNair & Lorr (1964) and Lishman (1972). Wilson-Barnett (1977) had found the MACL sufficiently sensitive to detect change in mood over time. However, in a study where the objective was to measure nurses' perceptions of their patients feelings, a major problem is that of validity of response. The effect of response bias in this study is unknown. It is not known how veridical a record the MACL was of the patients actual feelings. In other words whether or not patients were choosing not to disclose how they felt.

Although there can be no absolute criteria of external validity of a mood state or feeling, the factor analysis on the current data presented a similar profile to that of the previous work discussed. The failure of the adjectives describing feelings of self-deprocation to correspond to the factor Depression was present not only in the patient data but also in the nurse data. This did imply that Depression, in this instance, had achieved consensual understanding. This point does lead on to the differential use of language.

Although gross similarities had been observed in the use of language by the nurses and the patients through the factor analysis, it was likely that individual words were used in slightly dissimilar ways. It was not possible to correlate single patient and nurse factors because of the problem of double correlation. Items that had been found to correlate once within the factor analysis would have been correlated for a second time had a comparison of single factors been performed. That this was not done, however, did allow fine differences in the use of language to be reflected in the correlation that were observed.

The use of a different method had been considered. Johnston (1982) had employed lists of specific concerns or worries and inferred anxiety from their presence. It has already been discussed that this method is insensitive to questions of degree and it is specific to anxiety. It was not considered feasible to isolate antecedent conditions of Depression of Fatigue because what affects one person may not have the same affect on another. This is also true of inferring anxiety from patient concerns. The use of ranking adjectives or descriptors as a technique was also considered and it was considered to be too difficult a task to firstly discriminate between feelings experienced in rank order and secondly for another to judge this discrimination.

In sum, although there are apparent shortcomings with the method employed, the task of judging the feelings of another must be a task of approximation, a measure that allows an overall profile to be observed may be the most appropriate to employ.

The second major methodological issue in this section of the study was in relation to the measurement of judgement strategies. The nurses first completed the MACL as they thought that their patients had felt. They then completed it as they might have expected to feel had they been in that patients position, and lastly how they expected a typical patient to have felt. But the basic problem of eliciting a response as was done in this study may well have influenced the observations. The effect of contamination by order is unknown. The fixed order was maintained because of the effect of eliciting a stereotype on both these measures. But the observations of strategy use for the factors fatigue and depression were not possible to interpret for this reason. This part of the study can give directions for future work, however, where matched pairs of sub-groups of nurses are compared.

No account was taken of mathmatical relationships between the actual similarities between patients and nurses. This was not done because of the differences between the patients and nurses employed in the study. The patients were predominantly much older than the nurses and half were male. But there may have been a small degree of actual similarity present that was undetected. This must qualify the observations of the nurses' judgements of well-being of female patients to a point. But, there is a good case for including a dimension of similarity such as gender and observing any effects.

GENERAL CONCLUSION

The first half of this thesis concentrated upon the effect on judgement accuracy of adopting a particular strategy to start the judgement or decision process. That is, it examined the effect of the initial acquisition of information on further acquisition in relation to the accuracy of a clinical decision. There was an emphasis on mechanisms in relation to accuracy in the hope of identifying a strategy that could be taught and serve as an aid to decisions. The second half of the thesis aimed to simply describe the accuracy of another type of judgement. The exploration of mechanisms or strategy use was different in this case. The strategies identified for examination implied their use for the total process rather than simply initiating the process, and this was irrespective of their effect on accuracy which was not explored.

The work reported in the first half of the thesis did not demonstrate an advantage to performance of adopting a particular strategy in terms of the accuracy of a diagnosis. But, there was a consistent finding that the early use of general information that allowed a 'representative' context to be established led to the use of less information overall. The work reported in the second half of the thesis concentrated upon a clinical judgement of feelings or mood, a neglected area of study (Higgins, 1980). Physical feelings of tiredness or vigor were found to be judged more easily than those of anxiety or depression. The amount of anxiety or depression that patients experienced was overestimated by the nurses in an undifferentiated fashion. Further, awareness of physical feelings appeared to increase with experience. The actual strategy was unclear. There was some indication that less experienced nurses assumed patients to be similar to themselves, as did those who were relatively less cognitively complex. These exploratory findings can only be tentative and could be further explored in future research.

The work reported has highlighted some of the many methodological difficulties present. Where the task was to diagnose a post-operative complication, the method employed made that task too easy to allow effective analysis of the relationship of strategy use and accurate diagnosis. The exploration of strategy use in the second half of the thesis had problems which have been discussed and there was a failure to include other potential strategies that the nurse could have been employing rather than the two identified from the literature. The use of strategies could have been explored in both of the situations studied by the use of a common method ie. self-report protocols. This is where subjects give a verbal report of the information that they are selecting and the reason for that selection, but this method has the problem of the effects that are induced and influence behaviour by introspection and self-report.

Overall, there was an initial concentration on mechanisms with the intention of addressing accuracy in the first half of the thesis. The emphasis in the second half concentrated on accuracy with an initial exploration of mechanisms, although the two were not directly related as part of the study.

It was suggested in the introduction that the issues surrounding clinical judgement are important for nursing. A necessary pre-requisite for a nurse effectively caring for a patient is that he/she accurately perceives that patient's current need, and this has implications for a 'nursing process' or problem solving approach to patient care (McGilloway, 1980). This approach to the organisation of care pre-supposes that nurses can identify patient problems, both of a physical and psychological nature. The first part of the thesis reported a difficulty of the experimental task being too easy. This, in turn, suggested that this group of student nurses were well able to diagnose a physical complication such as a post-operative complication. The aim of the initial work reported was to identify a means of improving the performance of student nurses by teaching a strategy that characterised the performance of experienced nurses and doctors (Gordon, 1980; Elstein et. al., 1978). The strategy was that of early hypothesis formation that serves to direct search by delineating an end-point/goal and relevant attributes for testing. But, rather than providing explicit instructions to adopt a specific strategy (e.g. through the use of the task itself), the teaching intervention comprised an account of a hypothesised memory organisation. This could only have served to confuse rather than provide direct instructions about how to procede and as a consequence could not have had any effect.

The second part of the work shared the common theme of an ability by nurses to identify the physical condition of patients. There was an overall ability to judge both how tired patients felt and how vigorous they reported feeling. The correlations observed, although statistically significant, were of a low order ($r = 0.342$ and $r = 0.252$, respectively). This was suggested to reflect the difficulty of judging the experience of another through the use of a rating task. However, there is a common theme of awareness of physical condition that runs throughout this work. But, there was no awareness of psychological feelings. As has been discussed, this was in line with the findings of Johnston (1982). The problem-solving studies of Newell & Simon (1972) and the clinical reasoning studies of Elstein et. al. (1978) suggested that subjects employed a small number of heuristics or strategies. Further, a common and powerful heuristic adopted by subjects is to find analogies between the present problem and those experienced in the past. To extrapolate to the current case, it was evident that the nurses in this study when asked to rate the feelings of anxiety and depression experienced by their patients were unable to do so because of inexperience with a consideration of these feelings. There were no analogies to draw upon from past experience. The evidence of a generalised over-estimation of anxiety and depression suggests that the nurses were making a "best guess" (Gettys, Kelly & Peterson, 1973).

In terms of nursing the two pieces of work complement each other in that the first is concerned with a judgement of the physical condition of a hypothetical patient in a classroom setting, whereas the second is concerned with a judgement of both the physical and psychological condition of actual patients in a natural setting. It is of interest that judgement about physical condition was consistent throughout. Of further interest is the observation that first year nurses inferred more unhappiness and misery in patients than did their third year colleagues, even though this was in an undiscriminated form. This poses implications for the socialisation process within nursing in terms of reduced awareness of psychological factors as training and experience progress.

A consideration of the implications of the work for nursing directs itself mainly to nurse education. The type of problem-solving task employed in the first half can be adapted for a number of situations, i.e. different patient problems. It is suggested that this would be memorable, although it begs the question of effectiveness compared with other teaching methods. Real implications for nurse education lie in the province of teaching about psychological factors relating to hospital and illness. The work of McGuire et. al. (1974) and Moffic et. al. (1975) indicates the potential presence of psychopathology in a sizable minority of hospital patients. Nurses, because of their proximity to patients are in a position to be alerted and alert others to a possible need for expert intervention.

The considerable work documenting the presence of anxiety, its association with recovery and coping strategies (Johnston, 1980; Wilson-Barnett, 1977; Ridgeway & Mathews, 1982) suggests the importance of sustained anxiety as a patient problem. Yet these nurses were unaware of which patients were experiencing anxiety and to what extent. Nurses require teaching about the occurence of high levels of anxiety for certain patients, the issues that concern patients and the ways in which they can help their patients to cope. It was noted earlier that the most junior/inexperienced nurses inferred the most suffering and this does suggest desensisitisation as training progresses. It appears that there is a very real need for effective nurse teaching about psychological aspects of hospitalisation.

A summary of the main findings suggests that these nurses could judge the physical condition of patients in both a classroom and a natural setting, but had little awareness of which patients were experiencing negative affect.

APPENDIX 1

LECTURE

The aim of this talk is to tell you about the research that you are taking part in. I want to explain what I am doing and why, so that you can gain insight into some of the reasoning processes that may be involved in making a clinical judgement.

As nurses, we aim to be accurate in our clinical judgements. We aim to make an accurate diagnosis of the possibility of surgical complications in order to avail our patients of the medical care that they need, when they need it.

Currently, nurses are not taught how to use the relevant sign and symptom information.

This skill does not limit itself to surgical complications. It is necessary in any situation where we assess the well-being of a patient and plan the appropriate care, ie. a situation where we make a clinical judgement.

An example of a clinical judgement we all make frequently is the diagnosis of the experience of pain.

In order to improve this skill, it is necessary to attempt to understand what we are doing when making judgements of this nature.

This can be done by concentrating on the type of information that we need to look for and the best way to use it.

The first task that you did involved measuring the number of angles that you typically use to view people in your social world. The more angles from which you view a person, the more information you have at your disposal and the more accurate your impression is likely to be.

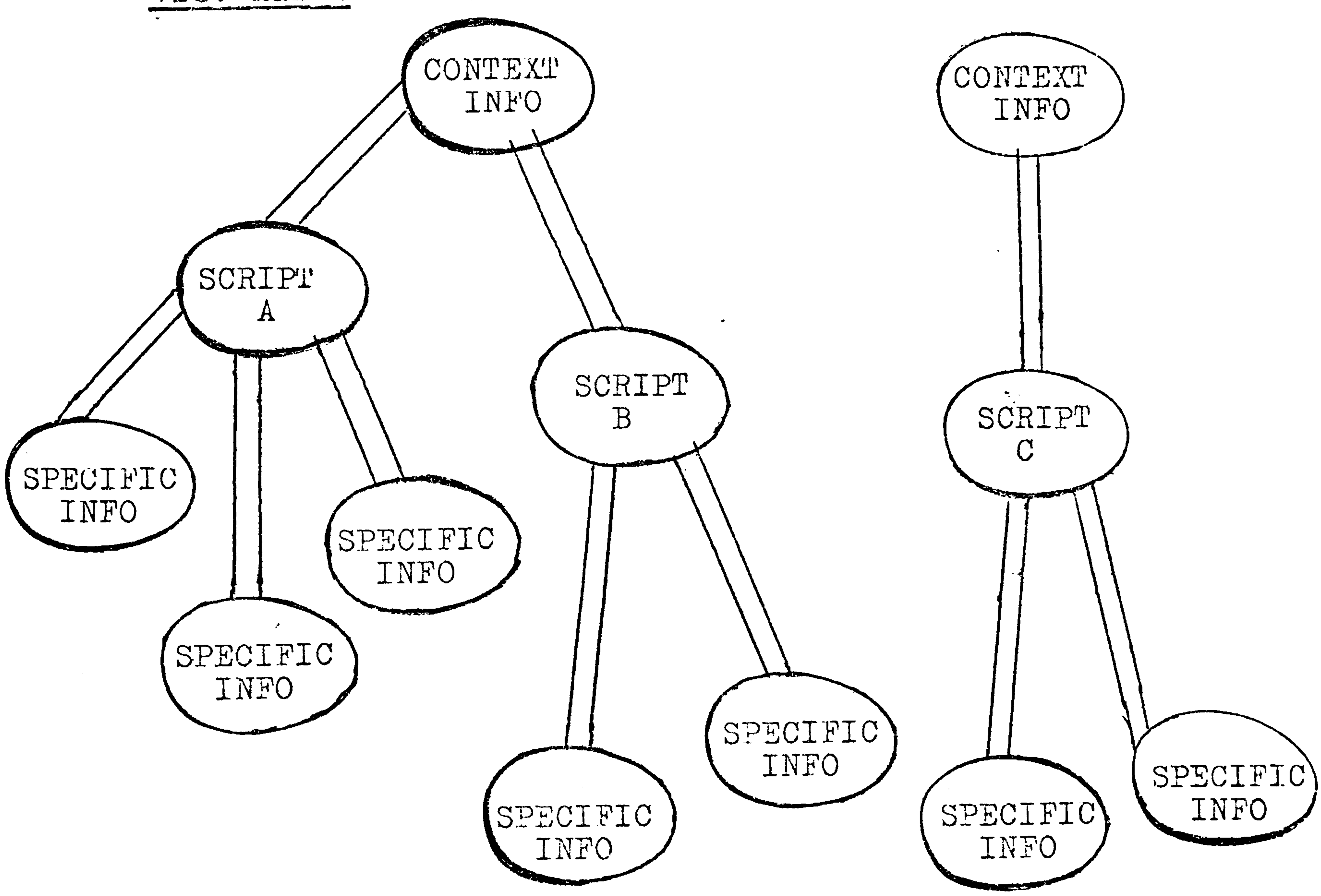
I will return to the relevance of this measure later. The insistence that no-one should write their name on any of the questionnaires is to ensure that they are completely anonymous. The entire purpose of this study is for you to get some benefit from it. There is no need for individual scores to be known.

I now want to talk to you about something that you may not have considered before. That is how human memory may operate and what sort of information will be of use to you when making a clinical judgement. Considering this information in relation to memory. The intention is to provide you with a strategy for making clinical judgements and a reason for using it.

Human memory can be described as a series of scripts or pictures. These scripts are the result of experiences that have been recorded and stored in memory.

Try to imagine your memory as a series of pictures that have been stores away. The pictures or scripts can be different types, they can be about an event like a conversation that you have had, or they can be a recorded memory of a clinical judgement that you have learned about. A clinical judgement such as a post-operative complication.

VIS. AID 1



This is a representation of 3 scripts, A, B and C. As you can see, stored along with the scripts are items of information. Items of context information and items of specific information.

Context information is general, relatively unchanging information. Specific information items are items that may change over relatively short periods.

These specific information items are facts that are associated with a particular script.

Whereas, the context information can be associated with one or more scripts.

The distinction between these two types of information will become clearer.

In order to retrieve or remember a script, the general, context information can serve as a clue or cue.

This will help you to remember one script or more than one.

To decide whether the script you are looking at is correct, or to distinguish between two scripts, it is necessary to check to see whether the specific information items associated with the script are present in the environment.

Context information can distinguish between the gross freatures of two scripts. But the specific facts of the remaining script need to be checked. It is not only necessary to look for facts to confirm a script. It is also necessary to test the script.

I want you to get used to the idea of testing the script that you are looking at.

Testing it by looking to see if there are specific facts about the patient that are not associated with the script and therefore indicate that the script is incorrect.

This may well seem rather abstract at present, and I will now illustrate all that I have been talking about in relation to the tasks that you have just performed.

The task that you have just done was designed to look at the strategies that people used when making clinical judgements that were based on the information that was available.

I shall be looking at the sort of information that people used first and if people, who were relatively more accurate, used different pieces of information or the same items in a different order.

The two types of information that I have just described were identified to be present in a clinical judgement situation by a study in the U.S.A.

<u>VIS. AID 2</u>

1. Information about features of the patient that are relatively stable, eg. sex, time since surgery, type of surgery etc.

   ie. context information

2. Information about the state of the patient that can change over short periods of time eg. pulse, blood pressure, respirations, how the patient feels etc.

   ie. state/specific information

The U.S. study found that the nurses they tested were very accurate in their clinical judgements.

They were people who had done a Masters Degree course, specialising in surgical nursing.

They used the first type of general or context information to set the scene and rule out an unlikely diagnosis.

They then used the second, more precise, patient state information, ie. specific information.

So, if you imagine dividing each of the tasks you have done in half, in the first half of the task these nurses set the scene, ie. retrieved the relevant scripts with context information.

In the second half, they examined these scripts using specific information items.

For an example, take the diagnosis of the experience of pain. It is likely that a patient who has had a relatively minor procedure, such as a cystoscopy, will experience some immediate post-operative discomfort on micturition.

The general, context information, that the patient has had a cystoscopy sets the scene.

You will expect this patient to experience less discomfort than the wound pain after abdominal surgery.

However, to obtain an accurate picture of exactly how much pain this patient is experiencing - we need to ask the patient.

It may be possible to alleviate this discomfort.

Another example is two patients who have had identical surgery are unlikely to experience identical discomfort. People do have very individual responses to surgery.

So, we can start with general information such as the knowledge that someone has had a specific operation.

This will lead us to expect them to experience a certain amount of pain, but the exact amount and the frequency with which to administer analgesia can only be gauged after asking the individual patient.

Asking for specific facts about that patient's current state that are not stable. People will be less accurate if they fail to use this specific information.

Another example that can be considered is the use of the general piece of context information - length of time since surgery.

This can be used to distinguish between two scripts that have been retrieved because of the information, that a patient has started to vomit after abdominal surgery. Four days after abdominal surgery, you would not expect a diagnosis of paralytic ileus for a patient who has recently started to vomit.

Therefore the context information does rule out an unlikely diagnosis for a patient who has started to vomit four days after surgery.

But you would need to know more specific details before you could distinguish between a gastrointestinal infection or an adverse reaction to a recently prescribed drug, eg. an antibiotic.

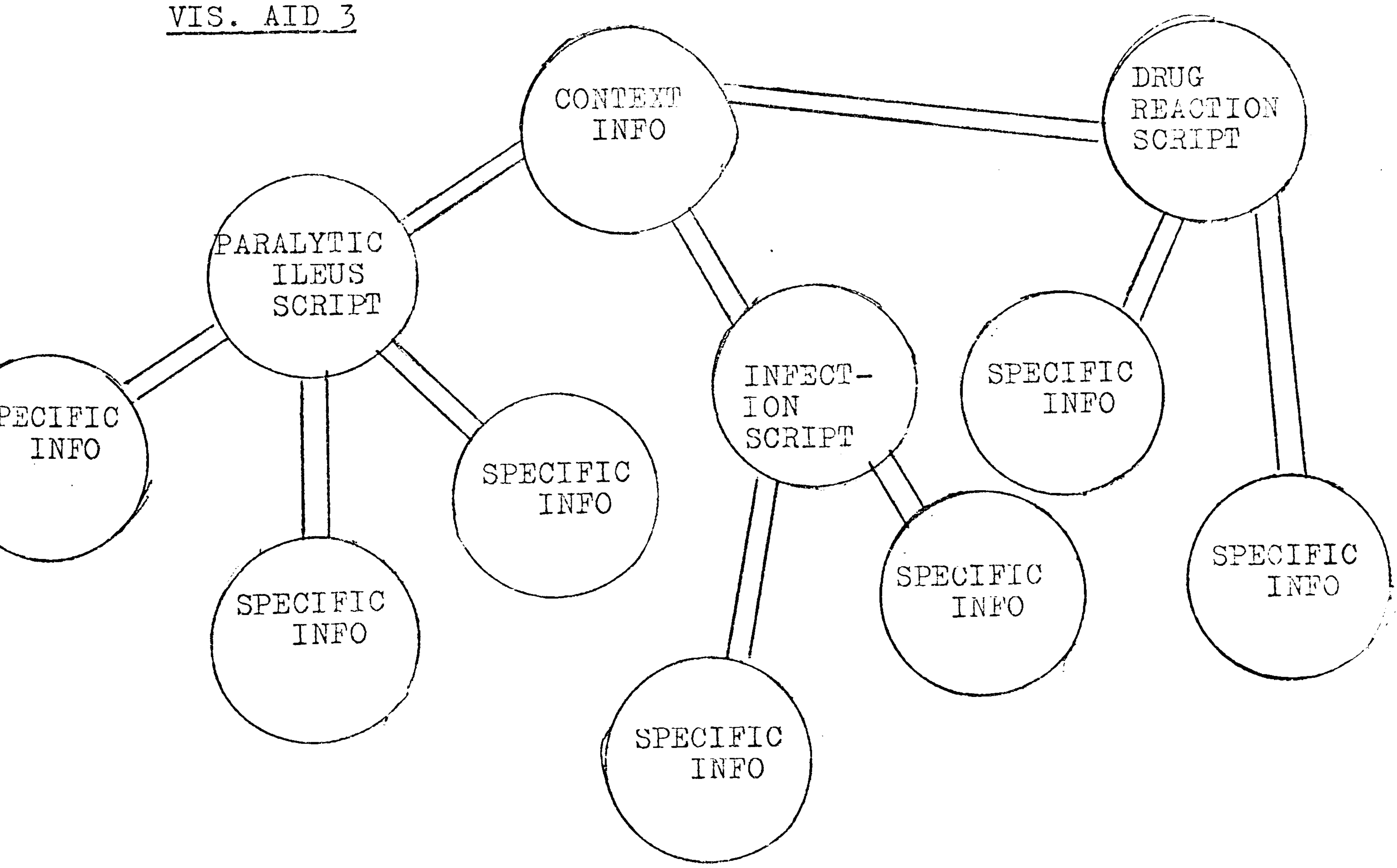
Information that could serve to reject one picture and confirm the other. Such as the patient's temperature, the pattern of drug prescription or laboratory reports.

Information checking the content of the script, ie. testing the scripts, rejecting one and accepting the one more likely to be correct.

Accuracy in the American study was associated with the initial use of general information, followed by the use of specific information relating to the current state of the patient.

Are there any questions so far?

VIS. AID 3



The context information, length of time since surgery, can rule out the paralytic ileus picture, because you know that surgery was four days' ago. But in order to distinguish between the other two, you need to know the specific signs that relate to the current patient state.

So you first form a general picture of the situation context using the general information.

Then you identify the appropriate script by looking at specific clues. To return to the relevance of the measure that assesses the number of angles people use to view their social world. Use of very few angles may indicate a reliance on the use of stereotypes of patients. A stereotype in this context being "an expectation for a patient to feel, respond and behave in a particular manner. Based on previous experience of people who appeared to be similar to that person or on nursing mythology". Take a simple and very obvious example of a patient that you know nothing about, other than that they have had a cholecystectomy. The myth or stereotype associated with people who suffer from cholecystitis is fair, fat, forty and female. However, cholecystitis is by no means confined to fair, fat females in middle-age. It occurs quite frequently in men and older people.

So this stereotype could be misleading if used as the sole basis for prediction in the absence of concrete evidence gained from personal knowledge of the person concerned or their records. The informational relevance of stereotypes is that they can be seen as scripts that are surrounded with general information from which we make predictions.

They can also be seen as a source of context information. The stereotype script also has specific information items encoded with it - these are the expectations that we have for that patient that I referred to earlier.

A stereotype can serve as a basis from which to start processing. We all <u>need</u> a basis to start.

If we had no expectations of how a person may respond after abdominal surgery, life would be impossible.

We do have an idea of a typical patient after a particular operation. Stereotype scripts are used to start processing, they set the scene and suggest the specific information items that may be present.

However, to obtain an accurate clinical picture we need to try and disprove the stereotype. Test it by looking for specific facts about the patient that are not encoded with our stereotype script.

What I am saying is that we need to go beyond the facts that the stereotypical picture lead us to expect, and look at other specific facts about the current state of the patient, in relation to the diagnosis of the experience of pain.

We have the stereotypical expectations, referred to earlier, in relation to types of surgery.

What is specific and is not associated with the stereotypical script is what the patient says about the experience, pulse rate and other signs of distress.

Not only do we have the useful expectations associated with different types of surgery, but nurses do have different expectations of the experience of pain for men and women.

Both in how much members of each group experience pain and in how much we think that they should express it.

Men are expected to be more stoic. Yet there is no conclusive evidence that demonstrates a difference in pain perception between the sexes. So we need to ask the patient and look for specific signs. Another possibility in relation to stereotypes is not just the obvious one that we have considered of their limited information content, but the possibility that we rarely know that we are using them. The more often that an item is retrieved from memory, the easier it becomes to retrieve.

Research in memory has shown this effect with items that are frequently retrieved.

Words that occur frequently in the English language, ie. common, everyday words, are easier to recall in a test of memory.

In these tests, people are typically shown a list of words, then a short time later asked to remember as many as possible. It does seem likely that we improve with practise when retrieving items from memory.

The fact that we do not recognise that we are using a stereotype may be the reverse side of the same coin.

In a simple recognition task where people look at a list of words that was presented earlier, and recognise as many as they can rather than writing them down without the original list, the rare English words are easier to recognise than the common ones.

So stereotypes are often used to start processing. They pop out easily, but because they are so familiar, we do not recognise that we are using them. They serve to start processing but we must look for specific facts to disconfirm our expectations, not just items to confirm them.

In order to be as accurate as possible, we need to use the available information in a way that promotes more accurate reasoning.

<u>VIS. AID</u>

1. Realise that a general picture is obtained to start with, using general, relatively unchanging context information. This may lead to a stereotypical picture and if we stop there we may make a mistake.

2. Look at the state of the individual patient, the current facts that change with the patient's state.

APPENDIX 2

PERSONAL DETAILS

PLEASE DO NOT WRITE YOUR NAME ON ANY OF THE TEST MATERIALS

The following details are necessary for an analysis of people's results and I would be grateful if you could supply the following:-

AGE:-

SEX:-

NUMBER OF 'A' LEVELS:-

NUMBER OF 'A' LEVELS:-

DEGREE, SUBJECT & CLASS:-

SECTION 1

In this section you are presented with 10 roles that will fit different people you know, the roles are:-

1. Yourself
2. A person that you dislike
3. Mother
4. A patient you would like to help
5. Father
6. Friend of the same sex
7. Friend of the opposite sex (or spouse)
8. Patient with whom you feel most uncomfortable
9. Ward Sister
10. Patient difficult to understand

For each role please select a different person that you know and pencil their initials beside the role for which you have been selected. These initials can be erased after you have finished to maintain complete anonymity.

Your task is now to rate on each of the scales below, there are no wrong or right answers. Just circle the number that you think best describes the person.

EXAMPLE

1.	Yourself	Outgoing	+3	+2	+1	-1	-2	-3	shy
		Adjusted	+3	+2	+1	-1	-2	-3	maladjusted

You may see yourself as being moderately shy so you have put a circle around -2 and you also see yourself as being well-adjusted so you have out a circle around +3. That was an example, now please rate each person on the scales as you think that the scale terms best describe that person.

1. Yourself

Outgoing	+3	+2	+1	-1	-2	-3	Shy
Adjusted	+3	+2	+1	-1	-2	-3	Maladjusted
Decisive	+3	+2	+1	-1	-2	-3	Indecisive
Calm	+3	+2	+1	-1	-2	-3	Excitable
Interested in others	+3	+2	+1	-1	-2	-3	Self-absored
Cheerful	+3	+2	+1	-1	-2	-3	Ill-humoured
Responsible	+3	+2	+1	-1	-2	-3	Irresponsible
Considerate	+3	+2	+1	-1	-2	-3	Inconsiderate
Independent	+3	+2	+1	-1	-2	-3	Dependent
Interesting	+3	+2	+1	-1	-2	-3	Dull

Appendix 3

Question: What or how much?

Card		Order	Answer
1	Cigarettes smoked daily		
2	Pain now		
3	Analgesia received		
4	Nausea/vomiting		
5	Anti-emetic		
6	Type of admission		
7	Type of surgery		
8	Time since surgery		
9	Time in theatre		
10	Level of consciousness		
11	Skin appearance		
12	Skin to touch		
13	Muscle tone		
14	Facial expression		
15	Level of anxiety/ distress		
16	B/P Pre-op		
17	B/P Current		
18	Pulse Pre-op		
19	Pulse current		
20	Temp. Pre-op		
21	Temp. current		
22	Resp. Pre-op		
23	Resp. current		
24	Cyanosis		

Card		Order	Answer
25	Expectoration		
26	Age		
27	Gender		
28	Weight		
29	I.V.I		
30	Naso-gastric tube		
31	Fluid chart		
32	Oral fluids currently taken		
33	Passed urine - bladder		
34	Bowel sounds		
35	Appearance of abdomen		
36	Dressing/ drainage		
37	Patients mobility		
38	Emotional & psychological History		
39	Reason for surgery Disease state		

Appendix 4

Question: What or how much? Paralytic Ileus

Card		Order	Answer
1	Cig. smoked daily		Does not smoke
2	Pain now		Abdominal discomfort
3	Analgesia rec.		Peth. 100mgs x 2 since op
4	Nausea/vomiting		x 6 since op
5	Anti-emetic		Maxalon x 2, nil effect
6	Type of admission		Emergency
7	Type of surgery		Appendicectomy
8	Time since surgery		3 days
9	Time in theatre		1 hour
10	Level of Consc.		Alert
11	Skin appearance		Pallor
12	Skin to touch		Mod. dry
13	Muscle tone		Rigid abdomen
14	Facial expression		Mod. distress
15	Level of anxiety/ distress		Mod. distress
16	B/P Pre-op		120/80
17	B/P current		120/80
18	Pulse Pre-op		80
19	Pulse current		90
20	Temp. Pre-op		TN
21	Temp. current		TN
22	Resp. Pre-op		24
23	Resp. current		30
24	Cyanosis		None
25	Expectoration		None

Card		Order	Answer
26	Age		24 years
27	Gender		Male
28	Weight		11 st
29	I.V.I		D/S + kve
30	Naso-gastric tube		Nil
31	Fluid chart		Yes
32	Oral fluids currently taken		30 mls
33	Passed urine - bladder		400 mls
34	Bowel sounds		Not present
35	App. of abdomen		Distension
36	Dressing/ Drainage		Nil
37	Patients mobility previously		Full
38	Emotional & psychological history		Nil
39	Reason for Surg./ disease state		Appendicectomy

Diagnosis:

Question: What or how much? Pulmonary Embolism

Card		Order	Answer
1	Cig. smoked daily		40+ daily
2	Pain now		Unilat. chest
3	Analgesia Rec.		100mgs Peth 1 hr ago
4	Nausea/vomiting		Nil
5	Anti-emetic		Nil
6	Type of admission		Emergency
7	Type of surgery		Appendicectomy
8	Time since surgery		24 hours
9	Time in theatre		1 hour
10	Level of consc.		Mod. alert
11	Skin appearance		Pallor
12	Skin to touch		Cold and clammy
13	Muscle tone		Normal
14	Facial expression		Distress
15	Level of anxiety/ distress		Distress
16	B/P Pre-op		140/80
17	B/P current		130/80
18	Pulse Pre-op		78
19	Pulse current		94
20	Temp. Pre-op		TN
21	Temp. current		35
22	Resp. Pre-op		24-28
23	Resp. current		32, dyspnoea
24	Cyanosis		Cyanosis
25	Expectoration		Frothy haemoptysis

Card		Order	Answer
26	Age		40
27	Gender		Female
28	Weight		11 st
29	I.V.I		12 hours post-op
30	Naso-gastric tube		Nil
31	Fluid chart		Yes
32	Oral fluids currently taken		Free fluids
33	Passed urine - bladder		400 mls
34	Bowel sounds		Present
35	App. of abdomen		NAD
36	Dressing/drainage		NAD
37	Patients mobility previously		Full
38	Emotional & psychological history		Nil
39	Reason for Surg./ disease state		Appendicitis

Diagnosis:

Question: What or how much? Urinary Retention

Card		Order	Answer
1	Cig. smoked daily		None
2	Pain now		Abdominal
3	Analgesia rec.		Peth 100 x 2
4	Nausea/vomiting		Nil
5	Anti-emetic		Nil
6	Type of admission		Emergency
7	Type of surgery		Appendicectomy
8	Time since surgery		16 hours
9	Time in theatre		1 hour
10	Level of consc.		Alert
11	Skin appearance		NAD
12	Skin to touch		Slightly moist
13	Muscle tone		Tense
14	Facial expression		Mod. distress
15	Level of anxiety/ distress		Some distress
16	B/P Pre-op		130/70
17	B/P current		130/70
18	Pulse Pre-op		82
19	Pulse current		90
20	Temp. Pre-op		35
21	Temp. current		34.8
22	Resp. Pre-op		24
23	Resp. current		30
24	Cyanosis		Nil
25	Expectoration		Nil

Card		Order	Answer
26	Age		60
27	Gender		Male
28	Weight		12 st
29	I.V.I		D/S + kve
30	Naso-Gastric tube		Nil
31	Fluid chart		Yes
32	Oral fluids currently taken		30mls hourly
33	Passed urine - bladder		HNPU
34	Bowel sounds		Present
35	App. of abdomen		Bladder palpable
36	Dressing/drainage		NAD
37	Patients mobility previously		Full
38	Emotional & psychological history		NAD
39	Reason for surg./ disease state		Appendicitis

APPENDIX 5

POSSIBLE PATIENT STATES

1. Peritonitis
2. No complication
3. Pulmonary embolism
4. Haemorrhagic shock
5. Chest infection
6. Paralytic ileus
7. Urinary retention
8. Thrombo-phlebitis
9. Chest infection

APPENDIX 6

Introduction and General Instructions

My name is Sheila Openshaw; I am a Nursing Researcher from Chelsea College. Mr O'Mullen is helping me with my current research project. The session and the tasks that you are about to do are part of that research project. This project is investigating the ways in which nurses make clinical judgements or diagnoses. It is also investigating reasoning strategies that can be taught to improve this skill. I shall be very grateful if you will help me by acting as subjects. The intention is that participation will be of practical use to you and I will explain it all to you after you have done the work.

I will now hand out a Questionnaire that I would like you to complete as quickly as possible. After this, I would like you to do two short tasks and I will explain what is involved after the Questionnaires have been completed. Each of you has a number, so please do not write your names on any of the sheets.

Please start by - reading the instructions and then fill out the Questionnaire and the personal details asked for. Please do not stop to think for too long about any item - do the task as quickly as possible. What is of interest is the first answer that you think of. There are no right or wrong answers to these sorts of questions. If there are any difficulties, I will come and explain.

APPENDIX 7

Diagnostic Task Instructions

Now I want you to do the two tasks and I will tell you what is involved. I want you to imagine that you are caring for a post-operative general patient. Your task is to judge the current condition of this patient. I want you to do this twice on each occasion, the patient's condition will be different. I will now give you a list of the possible conditions, between which you must decide. Plus two packs of cards and two lists (hand out - stating 1st and 2nd pack). Please do not touch anything yet. What I want you to do is to take your 1st list and pack. Do not look at the pack yet, just your list. Please do not at any point from now on look to see what your neighbour is doing. It will help me more if people do these tasks entirely alone.

As you can see, your lists contain a series of questions about information that you may wish to know about your patient. What I want you to do is to decide which question you wish to ask first. You can start anywhere in the list - there is no order - do not decide on No.1 unless you do really want that answer. There is no logic in the order.

Now, each question has a number beside it. Take the card with that number on it and only the card with that number on it - turn it over and you have the answer to your question. Now you all have your 1st answers. Write that down on your list beside the question that you asked, and put a No.1 in the column labelled order. It is important for me to know the order in which you ask each question, so write '2' for the second question, '3' for the third etc. Then, in the final space write down quickly the reason that you wanted to know that answer - very quickly and briefly. Now ask a second question and look at each card in turn, as you ask your question. Please do not look at any card that you are not specifically interested in. Ask as many questions as you need to, to assess the current condition of your patient. Not all the questions will be relevant - you will need to ask fewer than one-third. So start now by asking your questions, one by one. Record the answers, the order used and your reason for asking. So you will decide on Question 2 - write a '2' beside it - and look at the appropriate card. Write down the answer and your reason for asking. Then choose Question 3, and so on, until you have finished. Then write down your clinical judgement in the space provided at the bottom of your list. When you have done this, I want you to do the second task in an identical manner. Each task should take about five minutes each - please try to be quick, there is a lot to get through.

APPENDIX 8

CHI SQUARE COMPARING SUBJECTS USING $\geqslant$12 ITEMS WITH SUBJECTS USING $<$12 ITEMS, FOR THE GENERAL ITEMS, TYPE OF SURGERY AND LENGTH OF TIME SINCE SURGERY

2nd Task

2 x 2 contingency table derived from $\geqslant$ Ss group N. 11 and $<$ 12 Ss group N. 8.

Therefore the three most commonly occurring scores were deleted from $\geqslant$12 group.

	TYPE	TIME	TOTAL
12	A 14 E.F. = 8.75	B 12 E.F. = 17.24	26
12	C 18 E.F. = 23.2	D 51 E.F. = 45.7	69
	32	63	95

E.F. = expected frequency

$$X^2 = \frac{N(AD-BC/-\frac{N}{2})^2}{(A+B)\ (C+D)\ B+D)} \qquad \text{d.f.1}$$

$= 5.32$

From tables with d.f.1, the table value of $p<0.05$ is 3.84, the observed value 5.32 exceeds the table value. Therefore the Null hypothesis can be rejected at the $p<0.05$ level of significance.

APPENDIX 9

CHI SQUARE COMPARING SUBJECTS USING $\geqslant$12 ITEMS WITH SUBJECTS USING $\lessdot$12 ITEMS FOR GENERAL ITEMS, TYPE OF SURGERY AND LENGTH OF TIME SINCE SURGERY

3rd Task

2 x 2 contingency table derived from raw data of $\geqslant$Ss n.12 and $\lessdot$12 Ss n.8, the n of scores was equalised by delection of the most commonly occurring in group $\geqslant$12, ie.

$\geqslant$12 type surgery $\cancel{2}$,1,2,1,1,1,2,1,$\cancel{1}$,$\cancel{1}$,$\cancel{1}$
n.12 time surgery 3,$\cancel{2}$,3,$\cancel{2}$,2,1,$\cancel{2}$,2,$\cancel{2}$,10,$\cancel{3}$

$\lessdot$12 type surgery 3,3,4,5,1,1,4
n.8 time surgery 5,10,6,10,3,5

	TYPE	TIME	TOTAL
$\geqslant$12	A 9 E.F. = 10	B 21 E.F. = 20	30
$\lessdot$12	C 21 E.F. = 20	D 39 E.F. = 40	60
	30	60	90

E.F. = expected frequency

$$X^2 = \frac{N\ (1AD-BC1-\frac{N}{2})}{(A+B)(C+D)(A+C)(B+D)} \qquad d.f = 1$$

$$= .51$$

The observed value X^2 .51 fails to exceed the table value of 3.84 1df at $p \lessdot 0.05$, therefore the Null hypothesis is accepted.

APPENDIX 10

COMPARISON OF USE OF GENERAL INFORMATION IN THE 2nd HALF OF THE TASK, COMPARING SUBJECTS WHO WERE ACCURATE WITH THOSE WHO WERE INACCURATE

'INSTAT package on HARRIS computer

T-TEST between two independent groups

TASK 1

	N	MEAN		EQUAL VARIANCES	
			T.	DF.	SIG.
WRONG	12	16.43			
RIGHT	17	14.93	0.30	27	.763

TASK 2

	N	MEAN		EQUAL VARIANCES	
			T.	DF.	SIG.
RIGHT	20	10.41			
WRONG	9	17.42	1.31	27	.203

TASK 3

	N	MEAN		EQUAL VARIANCES	
			T.	DF.	SIG.
RIGHT	16	15.34			
WRONG	12	6.93	1.48	26	.152

APPENDIX 11

LECTURE

The aim of this talk is to tell you about the research that you are taking part in. I want to explain what I am doing and why, so that you can gain insight into some of the reasoning processes that may be involved in making a clinical judgement.

As nurses we aim to be accurate in all clinical judgements. We aim to make an accurate diagnosis of the possibility of surgical complications in order to avail all patients of the medical care that they need, when they need it.

Currently nurses are not taught how to use the relevant sign and symptom information.

This skill does not limit itself to surgical complications. It is necessary in any situation where we assess the well-being of a patient and plan the appropriate care ie. a situation where we make a clinical judgement.

An example of a clinical judgement we all make frequently is the diagnosis of the experience of pain.

In order to improve this skill, it is necessary to attempt to understand what we are doing, when making judgements of this nature.

This can be done by concentrating on the type of information that we need to look for and the best way to use it.

The first task that you did involved measuring the number of angles that you typically use to view people in your social world. The more angles from which you view a person, the more information you have at your disposal and the more accurate your impression is likely to be.

The insistence that no-one should write their name on any of the questionnaires is to ensure that they are completely anonymous. The entire purpose of this study is for you to get some benefit from it, there is no need for individual scores to be known.

I now want to talk to you about something that you may not have considered before. That is how human memory may operate and what sort of information will be of use to you when making a clinical judgement. Consider this information in relation to memory. The intention is to provide you with a strategy for making clinical judgements and reason for using it.

Human memory can be described as a series of scripts or pictures.

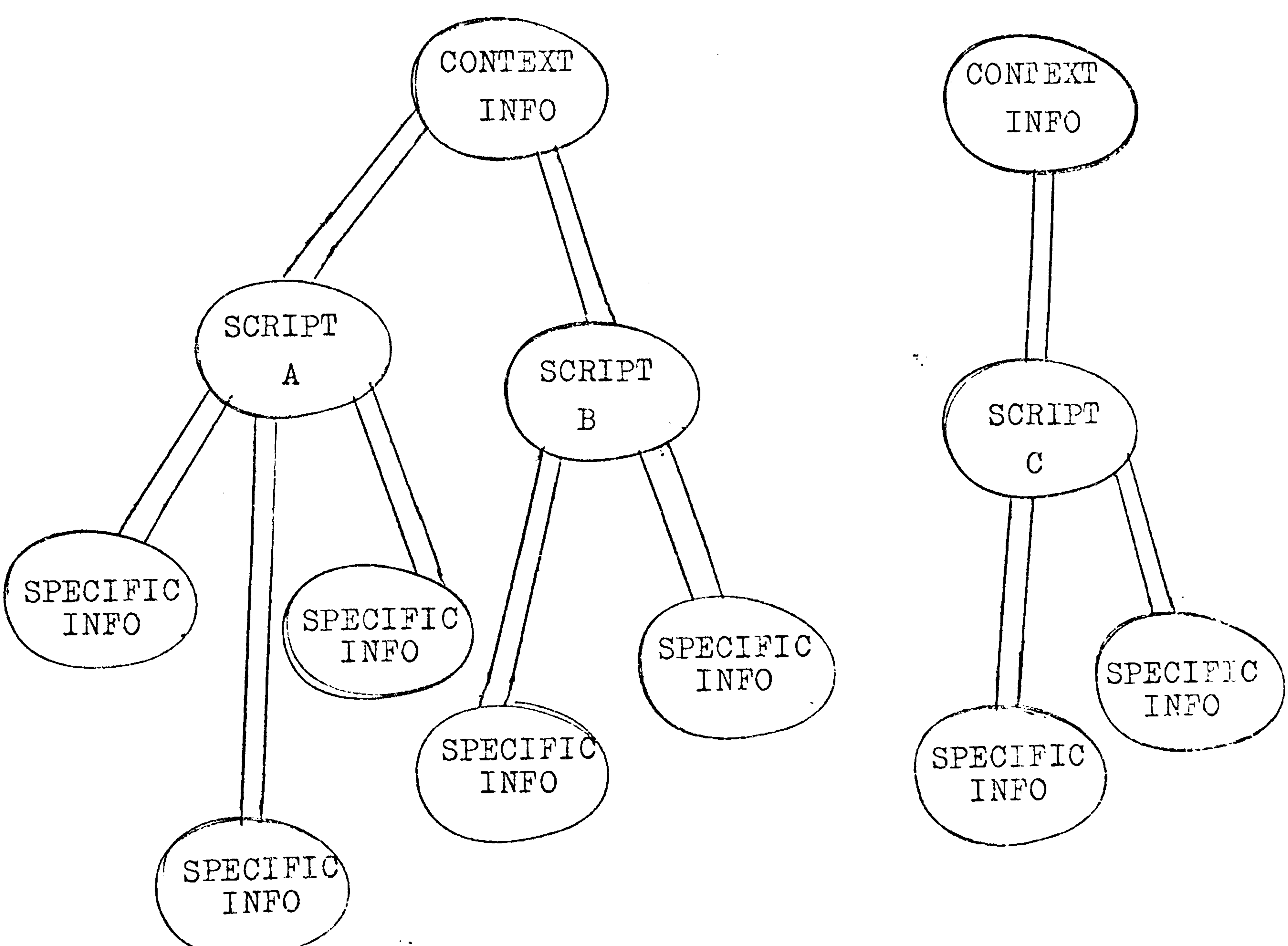
These scripts are the result of experiences that have been recorded and stored in memory.

Try and imagine your memory as a series of pictures that have been stored away.

The pictures or scripts can be of different types, they can be about an event like a conversation that you have had, or they can be a recorded memory of a clinical judgement that you have learned about.

A clinical judgement such as a postoperative complication.

VIS. AID 1



This is a representation of three scripts A, B and C.

As you can see, stored along with the scripts are items of information.

Items of context information and items of specific information.

Context information is general, relatively unchanging information.

Specific information items, are items that may change over relatively short periods.

These specific information items are facts that are associated with a particular script.

Whereas, the context information can be associated with one or more scripts.

The distinction between these two types of information will become clearer.

In order to retrieve or remember a script the general context information can serve as a clue or cue.

This will help you to remember one script or more than one.

To decide whether the script you are looking at is correct, or to distinguish between two scripts, it is necessary to check to see whether the specific information items associated with the script are present in the environment.

Context information can distinguish between the gross features of two scripts.

But the specific facts of the remaining script need to be checked.

It is not only necessary to look for facts to confirm a script

it is also necessary to test the script.

I want you to get used to the idea of testing the script that you are looking at.

Testing it by looking to see if there are specific facts about the patient that are not associated with the script and therefore indicate that the script is incorrect.

This may well seem rather abstract at present and I will now illustrate all that I have been talking about in relation to the tasks that you have just performed.

The task that you have just done was designed to look at the strategies that people used when making clinical judgements that were based on the information that was available.

I shall be looking at the sort of information that people used first, and if people who were relatively more accurate used different pieces of information or the same items in a different order.

The two types of information that I have just described were identified to be present in a clinical judgement situation by a study in the U.S.A.

VIS. AID 2

1. Info. about features of the patient that are relatively stable eg. age, sex, time since surgery, type of surgery etc.

   ie. Context information

2. Info. about the state of the patient that can change over short periods of time eg. pulse, B.P, respirations, how the patient feels etc.

   ie. State/specific info.

The U.S. study found that the nurses they tested were very accurate in their clinical judgements.

They were people who had done a Masters Degree course specialising in surgical nursing.

They used the first type of general or context information to set the scene and rule out an unlikely diagnosis.

Then they used the second, more precise patient state information ie. specific information.

So, if you imagine dividing each of the tasks you have done in half,

in the first half of the task these nurses set the scene ie. retrieved the relevant scripts with context information.

In the second half they examined these scripts using specific information items.

For an example take the diagnosis of the experience of pain.

It is likely that a patient who has had a relatively minor procedure

The general, context information, that the patient has had a cystoscopy sets the scene.

You will expect this patient to experience less discomfort than the wound pain after abdominal surgery.

However, to obtain an accurate picture of exactly how much pain this patient is experiencing we need to ask the patient.

It may be possible to alleviate this discomfort.

Another example is two patients who have had identical surgery,

are unlikely to experience identical discomfort.

People do have very individual responses to surgery.

So we can start with general information such as the knowledge that someone has had a specific operation.

This will lead us to expect them to experience a certain amount of pain,

but the exact amount and the frequency with which to administer analgesia can only be guaged after asking the individual patient.

Asking for specific facts about that patient's current state that are not stable.

People will be less accurate if they fail to use this specific information.

Another example that can be considered is the use of the general piece of context information - length of time since surgery.

This can be used to distinguish between two scripts that have been retrieved because of the information that a patient has started to vomit after abdominal surgery.

Four days after abdominal surgery you would not expect a diagnosis of an adverse reaction to anaesthesia for a patient who has recently started to vomit.

Therefore the context information does rule out an unlikely diagnosis for a patient who has started to vomit four days after surgery.

But you would need to know more specific details before you could distinguish between a gastrointestinal infection or an adverse reaction to a recently prescribed drug eg. an antibiotic.

Information that could serve to reject one picture and confirm the other such as the patient's temperature, the pattern of drug prescription or laboratory reports.

Information checking the content of the script ie. testing the scripts, rejecting one and accepting the one more likely to be correct.

Accuracy in the American study was associated with the initial use of general information. Followed by the use of specific information relating to the current state of the patient.

Are there any questions so far?

To return to the processes involved in memory and how the information items relate to them.

Human memory is a lot more complex than the model that I have described and no-one really knows the exact structure.

But one of the very important findings is that scripts or pictures are encoded with their context and the context can retrieve the corresponding script.

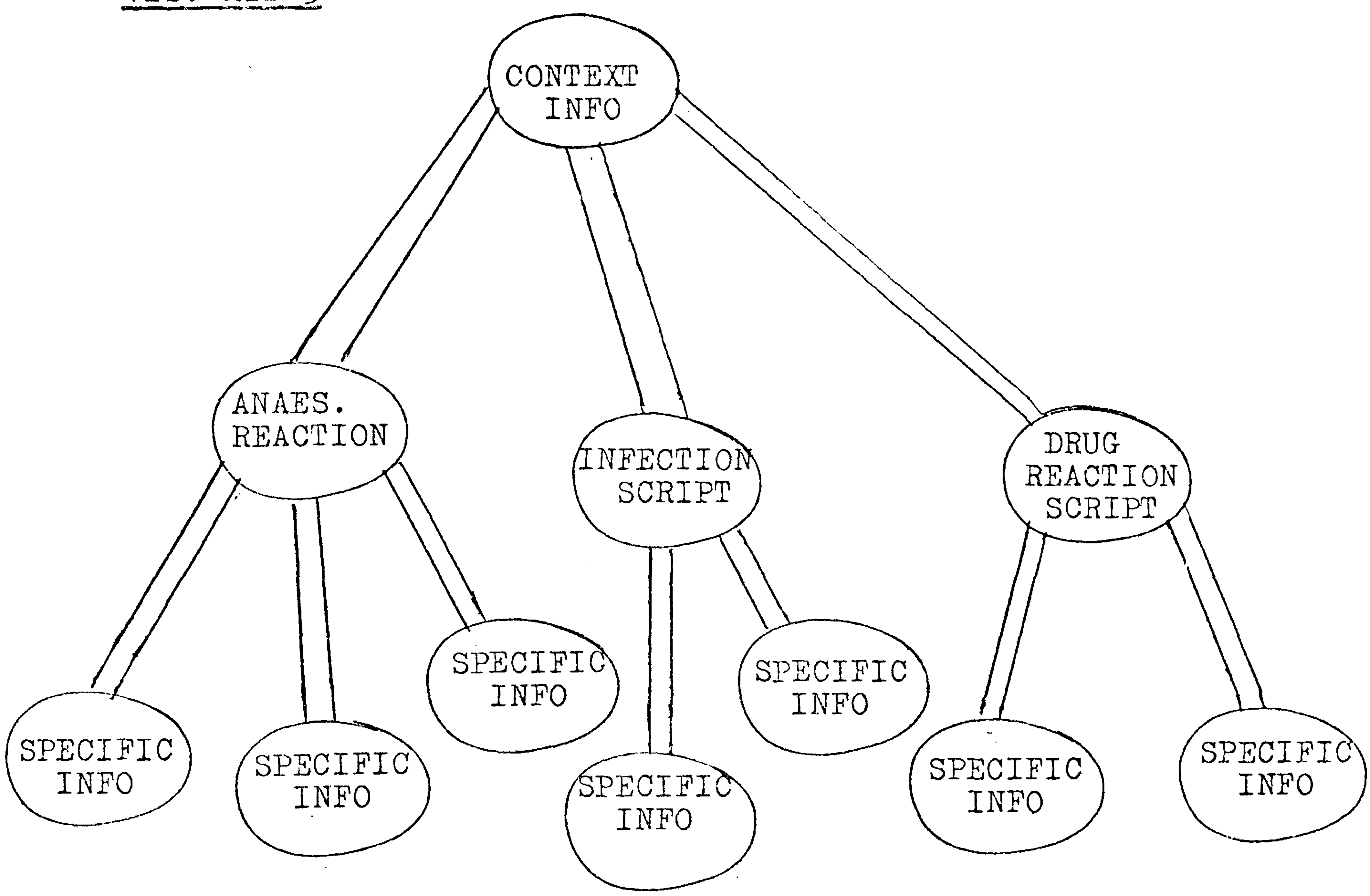
In the example of a patient vomiting after abdominal surgery.

The general information that the surgery was abdominal plus the abnormal patient state can lead to the retrieval of three possible diagnostic pictures.

Adverse anaesthetic reaction, a gastrointestinal infection and an adverse drug reaction.

All three being possible cause of vomiting after surgery.

VIS. AID 3



The context information, length of time since surgery, can rule out the anaesthetic reaction picture because you know that surgery was four days ago.

But in order to distinguish between the other two you need to know the specific signs that relate to the current patient state.

So you first form a general picture of the situation context using the general information.

Then you identify the appropriate script by looking at specific clues

VIS. AID

1. Realise that a general picture is obtained to start with, using general, relatively unchanging, context information. This may lead to a stereotypical picture and if we stop there we may make a mistake.

2. Look at the state of the individual patient, the current facts that change with the patient's state.

APPENDIX 12

POSSIBLE PATIENT STATES

1. HAEMORRHAGIC SHOCK
2. DEEP VEIN THROMBOSIS
3. SUBPHRENIC ABSCESS
4. ILIAC ABSCESS
5. NO COMPLICATION
6. BREAKDOWN OF ANASTOMOSIS
7. RENAL FAILURE
8. URINARY RENTENTION
9. BURST WOUND
10. REDUCED BLOOD VOLUME
11. ATELECTASIS
12. PULMONARY EMBOLISM
13. INHALATION OF VOMIT
14. ADHESIONS
15. PERITONITIS
16. PELVIC ABSCESS
17. WOUND INFECTION
18. CHEST INFECTION
19. PARALYTIC ILEUS
20. SEPTICAEMIA
21. THROMBOPHLEBITIS
22. HERNIA

APPENDIX 13

	TASK 1 (B_1)		TASK 2 (B_2)		TASK 3 (B_3)		TS	TA
	C_1	C_2	C_1	C_2	C_1	C_2		
	-	16.7	40	40	50	16.7	163.4	
	42.9	42.9	42.9	28.6	20	20	197.3	
	11.1	33.3	28.6	-	28.6	-	101.6	
	42.9	-	33.2	12.5	40	20	148.7	
	33.3	-	33.3	12.5	40	20	139.1	
	33.3	33.3	50	25	33.3	33.3	208.2	
	36.4	18.2	37.5	37.5	33.3	33.3	196.2	
	40	40	37.5	25	42.9	14.3	1997.7	2602.4
	20	10	75	-	16.7	16.7	138.4	
A_1	33.3	-	100	-	50	-	183.3	
	36.4	18.2	40	30	57.1	-	181.7	
	-	42.9	30	20	57.1	-	150	
	22.2	11.1	28.6	-	25	-	86.9	
	60	10	71.4	-	83.3	-	224.7	
	25	25	16.7	16.7	33.3	33.3	150	
	11.1	33.3	33.3	22.2	33.3	-	133.2	
	54.5	9.1	37.5	25	42.9	-	169	
	7.7	46.2	13.3	26.7	14.3	14.3	122.5	
	60	20	40	20	44.4	-	184.4	
	28.6	28.6	28.6	14.3	50	-	150.1	
	37.5	-	55.6	-	66.7	-	159.8	
	40	10	57.1	-	66.7	22.2	196	
	33.3	16.7	33.3	11.1	28.6	28.6	151.6	
A_2	11.1	22.2	40	10	42.9	14.3	140.5	2576.1
	14.3	42.9	42.9	14.3	30	-	144.4	
	50	-	33.3	22.2	57.1	-	162.6	
	42.9	-	37.5	25	57.1	-	162.5	
	50	20	60	-	50	-	170	
	57.1	-	50	-	60	-	157.1	
	14.3	57.1	50	16.7	50	25	213.1	
	28.6	28.6	57.1	-	42.9	-	157.2	
	25	25	9.1	9.1	28.6	28.6	125.4	
TC	TC 992.8	661.3	1343.4	464.4	1376.1	340.6		
TB	TB 1654.1		18078		1716.7			
TC^2								
VAR	276.1	260.7	328.4	147.8	246.2	155.7		

Variance derived from $\dfrac{\Sigma X^2 - (\Sigma X)^2 n}{n - 1}$

$\therefore$ F max $= \dfrac{328.4}{147.8} = 2.2$ when there are 6 columns and n-1 is 31 df (6.31). From Tables at $p < 0.05$ df (6.30) F $\doteq$ 3.81 $\therefore$ the observed value of 2.2 fails to reach significance and we can assume homogeneity of variance.

ABC SUMMARY TABLE

	B_1		B_2		B_3		TA
	C_1	C_2	C_1	C_2	C_1	C_2	
A_1	447.9	334.9	698.1	270	643.9	207.6	2602.4
A_2	544.9	326.4	645.3	194.4	732.2	133	2576.2
TA	992.8	661.3	1343.4	464.4	1376.1	340.6	5178.6

AB SUMMARY TABLE

	B_1	B_2	B_3	TA
A_1	782.8	968.1	851.5	2602.4
A_2	871.3	839.7	865.2	2576.2
TB	1654.1	1807.8	1716.7	5178.6

AC SUMMARY TABLE

	C_1	C_2	TA
A_1	1789.9	812.5	2602.4
A_2	1922.4	653.8	2576.2
TC	3712.3	1466.3	5178.6

BC SUMMARY TABLE

	C_1	C_2	TB
B_1	992.8	661.3	1654.1
B_2	1343.4	464.4	1807.8
B_3	1376.1	340.6	1716.7
TC	3712.3	1466.3	5178.6

A(a) = Cognitive Complexity $a = 2$
B(b) = Task $b = 3$
C(c) = Half of task $c = 2$
$n = 16$

$$1 = (\Sigma X)^2 = \frac{5178.6^2}{(16)(2)(3)(2)} = 139676.6$$

$$2 = \Sigma X^2 = 215363.71$$

$$3 = \frac{\Sigma Ai^2}{nbc} = \frac{(2602.4+2576.2^2)}{(16)(3)(2)} = 139680.1$$

$$4 = \frac{\Sigma Bj^2}{nac} = \frac{(1654.1^2+1807.8^2+1716.7^2)}{(16)(2)(2)} = 139863.2$$

$$5 = \frac{\Sigma Ck^2}{nab} = \frac{(3712.3^2+1466.3^2)}{(16)(2)(3)} = 165950.06$$

$$6 = \frac{\Sigma ABij^2}{nc} = \frac{(782.8^2+968.1^2+851.5^2+871.3^2+839.7^2+865.2^2)}{(16)(2)} = 140246.1$$

$$7 = \frac{\Sigma ACik^2}{ab} = \frac{(1789.9^2+812.5^2+1922.4^2+653.8^2)}{(16)(3)} = 166395.3$$

$$8 = \frac{\Sigma BCjk^2}{na} = \frac{(992.8^2+661.3^2+1343.4^2+464.4^2+1376.1^1+340.6^2)}{32} = 170406.9$$

$$9 = \frac{\Sigma ABCijk^2}{n} = \frac{(447.9^2+\ldots\ldots+133^2)}{16} = 171386.4$$

$$10 = \frac{\Sigma Ts^2}{n} = \frac{(163.4^2+\ldots\ldots+125.4^2)}{6} = 144804.6$$

$$11 = \frac{\Sigma TBs^2}{c} = \frac{(93168.2+112087.1+99738.1)}{2} = 152496.7$$

$$12 = \frac{\Sigma TCs^2}{c} = \frac{(429954.6+101249.3)}{3} = 177068.$$

ANOVA SUMMARY TABLE

Source of variance	df	Sums of Squares	Mean Squares	Variance Ratio	Probability
Between subjects	31	5127.9			
(A) Cognitive Comp.	15	3.5	0.233	.0006	n.s
Between groups error	30	5124.4	170.8		
Within subjects	160	70559.2			
(B) Tasks	2	186.6	93.3	.48	n.s
A x B Cog Com x Tasks	2	379.4	189.7	1.6	n.s
Within subjects error B	60	7126.2	118.77		
Half of task	1	26273.5	26273.5	142.09	$p < 0.001$
A x C Cog Comp x $\frac{1}{2}$ task	1	441.7	441.7	2.4	n.s
Within subjects error C	30	5548.3	184.9		
Task x $\frac{1}{2}$ task BC	2	4270.24	2135.12	4.9	$p < 0.025$
A x B x C	2	155.06	77.5	.177	n.s
Be within subjects error	60	26178.4	436.3		
TOTAL	191	75687.1			

APPENDIX 14

STUDENT'S t-TEST FOR INDEPENDENT MEANS

TASK I COMPARING GENERAL INFORMATION USE IN THE SECOND HALF OF THE TASK FOR ACCURATE AND INACCURATE SUBJECTS

x_a		x_b	
$n = 33$		$n = 15$	
1) Σx_a	$= 612.5$	Σx_b	$= 335.2$
2) Σx_a^2	$= 19588.01$	Σx_b^2	$= 10414.3$
3) $(\Sigma x_a)^2$	$= 375156.25$	$(\Sigma x_b)^2$	$= 112359.4$
4) n	$= 33$	n	$= 15$
5) $\frac{(\Sigma x_a)^2}{n}$	$= 11368.4$	$\frac{(\Sigma x_b)^2}{n}$	$= 7490.6$
6) $\bar{x}_a$	$= 18.6$	$\bar{x}_b$	$= 22.35$
7) $\Sigma x^2_a - \frac{(\Sigma x_a)^2}{n}$	$= 8219.61$	$\Sigma x^2_b - \frac{(\Sigma x_b)^2}{n}$	$= 2923.7$

$$\frac{\bar{x}_a - \bar{x}_b}{\sqrt{\dfrac{\Sigma x^2_a - \frac{(\Sigma x_a)^2}{n} + \Sigma x_b^2 - \frac{(\Sigma x_b)^2}{n}}{(n_a - 1) + (n_b - 1)} \times \dfrac{1}{n_a} + \dfrac{1}{n_b}}}$$

$$\frac{-3.75}{\sqrt{\dfrac{11143.31}{46} \times .09667}}$$

$$= -.774$$

Observed value < 1 $\therefore$ fails to reach all acceptable level of significance

APPENDIX 15

STUDENT'S t-TEST FOR INDEPENDENT MEANS

TASK II COMPARING CONTEXT INFORMATION USE IN THE SECOND HALF OF THE TASK FOR ACCURATE AND INACCURATE SUBJECTS

	xa	xb
	$n = 30$	$n = 18$
1)	$\Sigma xa^2 = 524.8$	$\Sigma xb = 244.5$
2)	$\Sigma xa^2 = 17438.62$	$\Sigma xb^2 = 5955.31$
3)	$(\Sigma xa)^2 = 275415.04$	$(\Sigma xb)^2 = 57780.25$
4)	$n = 30$	$n = 18$
5)	$(\Sigma xa)^2 = 9180.5$	$(\Sigma xb)^2 = 3321.1$
6)	$\bar{x}a = 17.5$	$\bar{x}b = 13.6$
7)	$\Sigma xa^2 - \dfrac{(\Sigma xa)^2}{n}$	$\Sigma xb^2 - \dfrac{(\Sigma xb)^2}{n}$
	$= 8258.12$	$= 2634.21$

$$\frac{\bar{x}a - \bar{x}b}{\sqrt{\dfrac{\Sigma xa^2 - \dfrac{(\Sigma xa)^2}{n} + \Sigma xb^2 - \dfrac{(\Sigma xb)^2}{n}}{(na - 1) + (nb - 1)} \times \dfrac{1}{na} + \dfrac{1}{nb}}}$$

$$\frac{3.9}{\sqrt{\dfrac{10892.33}{46} \times 0.0889}}$$

$$t = .85$$

APPENDIX 16

STUDENT'S t-TEST FOR INDEPENDENT MEANS

TASK III COMPARING CONTEXT INFORMATION USE IN THE SECOND HALF OF THE TASK FOR ACCURATE AND INACCURATE SUBJECTS

ACCURATE X_a	INACCURATE X_b
$\Sigma X_a = 438.3$	$\Sigma X_b = 75.4$
$\bar{X}_a = 11.2$	$\bar{X}_b = 8.4$
$n = 39$	$n = 9$

$$\bar{X}_a - \bar{X}_b = 2.8$$

The observed difference between the means renders the computation of a further t-test necessary as the computed t will be >1.

APPENDIX 19

2 x 2 ANOVA - UNEQUAL CELL FREQUENCIES - LEAST SQUARES SOLUTION. TASK I

CELLS

	1-3b	1-3=4+b2	4+b3
Correct Subjects a_1	14,14,16,20, 8,12, 9, 6,14,11, 14,19,12, 9	21,18,6,10 18,13,13,	26,13,21,23,20, 15,21,14,14, 27,15
Incorrect Subjects a_2	10,12,22,15	13,18,11	21,18,9,18, 19,25,23,13

CELL FREQUENCIES

	b1	b2	b3	TOTAL
a_1	n_{11} 14	n_{12} 7	n_{13} 11	n_{1j} 32
a_2	n_{21} 4	n_{22} 3	n_{23} 8	n_{2j} 15
	n_{ij} 18	n_{2j} 10	n_{3j} 19	47

CELL TOTALS

	b1	b2	b3	
a_1	AB_{11} 179	AB_{12} 99	AB_{13} 209	A_1 487
a_2	AB_{21} 59	AB_{22} 42	AB_{23} 146	A_2 247
	B_1 238	B_2 141	B_3 355	G 734

$$(1) = G^2/n \qquad = 11462.9$$

$$(2) = \quad ^2 \qquad = 12708$$

$$(3) = (Ai^2/ni) = \frac{487^2}{32} + \frac{247^2}{15} \qquad = 11478.8$$

$$(4) = (Bj^2/nj) = 238^2/18 + 141^2/10 + 355^2/19 \qquad = 11767.9$$

$$(5) = (ABij^2/nij) = \frac{179^2}{14} + \frac{99^2}{7} + \frac{209^2}{11} + \frac{59^2}{4} + \frac{42^2}{8} + \frac{146^2}{8}$$

$$= 11782.45$$

S.S. ab (adj)

CELL MEANS

a_1	12.8	14.14	19
a_2	14.75	14	18.25
dj	-1.95	.14	.75
wj	3.1	2.1	4.6 wj = 9.8
wjdj	1.15	.3	3.45 wjdj = 4.9

$$SS\ ab(ddj) = wjd^2j - \frac{(\Sigma wjdj)^2}{wj}$$

$$= 11.8 + .042 + 2.6 - \frac{(4.9)^2}{9.8}$$

$$= 14.4 - 2.45$$

$$= 11.98$$

SS cells = (5) - (1) = 319.55

SSa = (3) - (1) = 15.9

SSb = (4) - (1) = 305

The adjusted sums of squares

SSa (adj) = SS cells - SSab (adj) - SSb = 2.57

SSb (adj) = SS cells - SSab (adj) - SSa = 291.7

SS error = (2) - (5) = 925.55

SUMMARY TABLE

Source of variance	SS	df	MS	VR	P
A Correct/ Incorrect	2.57	1	2.57	.115	ns
B Question Use	291.7	2	145.53	6.4	$p<.01$
AB	11.98	2	5.99	.26	ns
Error	925.55	41	22.6		

The table value for $p<0.01$ df (2.40) is 5.18 ∴ an observed value of 6.4 df (2.41) can reject the Null hypothesis of no difference in question use over the three conditions of B at the $p<0.01$ level of significance.

APPENDIX 20

2 x 2 ANOVA - UNEQUAL CELL FREQUENCIES
LEAST SQUARE SOLUTION   TASK II

CELLS

	(1-3)b_1	(1-3=4+)b_2	(4+)b_3
Correct Subjects a_1	20,14,11,6,18, 14,9,17,8, 4,14,16,15 6,8,12,16	16,10,11,21,11	29,9,13, 15,18,19, 9,18
Incorrect Subjects a_2	19,13,7,11,19,19 19,14,14,15,9,13	8,15,5,17	24,15

CELL FREQUENCIES

	b_1		b_2		b_3		TOTAL	
a_1	n_{11}	17	n_{12}	5	n_{13}	8	n_{1i}	30
a_2	n_{21}	12	n_{22}	4	n_{23}	2	n_{2i}	18
	n_{ij}	29	n_{2j}	9	n_{3j}	10		48

CELL TOTALS

	b_1		b_2		b_3			
a_1	AB_{11}	208	AB_{12}	69	AB_{13}	130	A_1	407
a_2	AB_{21}	170	AB_{22}	45	AB_{23}	39	A_2	254
	B_1	378	B_2	114	B_3	169	G	661

$$(1) = G^2/n \qquad = 9102.5$$

$$(2) = \Sigma X^2 \qquad = 10327$$

$$(3) = \Sigma(A^2i/ni) = \frac{407^2}{30} + \frac{254^2}{18} \qquad = 9105.8$$

$$(4) = \Sigma(B^2j/nj) = \frac{378^2}{29} + \frac{114^2}{9} + \frac{169^2}{10} \qquad = 9227.13$$

$$(5) = \Sigma\Sigma(AB^2ij/nij) = \frac{208^2}{17} + \frac{69^2}{5} + \frac{130^2}{8} + \frac{170^2}{12} + \frac{45^2}{4} + \frac{39^2}{2}$$

$$= 9284.65$$

S.S. ab (adj)

CELL MEANS

	b_1	b_2	b_3
a_1	12.2	13.8	16.25
a_2	12	11.25	19.5
dj	.2	2.55	-3.25
wj	7.03	2.2	1.6
wjdj	1.4	5.61	-1.65

$\Sigma wj = 10.83$

$\Sigma wjdj = 5.36$

$$SS\ ab(adj) = \Sigma wjd^2) - \frac{(\Sigma wjdj)^2}{wj}$$

$$.28 + 14.3 + 16.9 - 2.7 = 28.8$$

SS cells = (5) - (1) = 182.2

SSa = (3) - (1) = 3.3

SSb = (4) - (1) = 124.6

The adjusted sums of squares

SSa (adj) = SS cells - SSab (adj) - SSb = 28.8

SSb (adj) = SS cells - SSab (adj) - SSa = 150.1

SS error = (2) - (5) = 1042.4

Source of variance	SS	df	MS	VR	P
A Correct/ Incorrect	28.8	1	28.8	1.2	ns
B Question Use	150.1	2	75.05	3.03	<.10 ns
AB	28.8	2	14.4	.58	ns
Error	1042.4	42	24.8		

The table value for <0.05 df (2.40) is 3.23, df (2.60) = 3.15 ∴ an observed value for A, B & AB do not allow rejection of the Null hypotheses in diagnoses.

APPENDIX 21

X^2 Correct Subjects

Difference in the number of subjects asking the general questions more in position 1-3 and those asking more later in the task as a function of the tasks I-III.

	TASK I	TASK II	TASK III	
1.3	(*18.4) 14	(*17.2) 17	(*22.4) 27	58
1-3 = 4+	(*6.7) 7	(*6.2) 5	(*8.1) 9	21
4+	(*6.97) 11	(*6.5) 8	(*8.5) 3	22
	32	30	39	101

*Expected frequencies

It can be observed that the expected frequencies of those subjects asking an equal number of questions in positions 1-3 and later in the task are less than one whole number different to the observed frequencies and have been dropped from the analysis.

Adjusted Table

	TASK I	TASK II	TASK III	
1-3	(*18.1) 14	(*18.1) 17	(*21.75) 27	58
4+	(*6.9) 11	(*6.9) 8	(*8.25) 3	22
	25	25	30	80

$$X^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(Oij - Eij)^2}{Eij}$$

$= 8.175$ df $(3-1)(2-1) = 2$

Table value df 2 at $p<0.025 = 7.38$ $\therefore$ with an observed value of 8.175 the Null hypothesis of no differences can be rejected

APPENDIX 22

FISHER EXACT PROBABILITY

Inaccurate subjects. The difference in the number of subjects asking more general questions in positions 1-3 and those asking more later in the task as a function of Tasks I & II.

	TASK I	TASK II	
1-3	4	12	16
4+	8	2	10
	12	14	26

In order to use significance levels in Siegel (1956) observed frequencies reduced by a constant proportion of $\frac{1}{2}$.

	TASK I	TASK II	
1-3	2	6	8
4+	4	1	5
	6	7	

From Tables

$$\left.\begin{array}{r} A + B = 8 \\ C + D = 5 \\ B = 6 \end{array}\right\} \quad B = 0 \text{ at } p < 0.05 \text{ level of significance}$$

$\therefore$ an observed value of B = 1 does not allow rejection of the Null hypothesis.

APPENDIX 23

FISHER EXACT PROBABILITY

Inaccurate subjects. The difference between the number of subjects asking more general questions in position 1-3 than later in the task as a function of Task II & III

	TASK I	TASK III	
1-3	4	8	12
4+	8	0	8
	12	8	

From Tables

$A + B = 12$ Table value of $D = 0$ and

$C + D = 8$ observed value of $D = 0$

$B = 6$ at $p < 0.025$

APPENDIX 24

Question: What or how much?

	Card Question	Order	Answer
1	Level of consciousness		
2	Disease State		
3	Appearance/Dressing		
4	Frame-size		
5	State of bladder		
6	Age		
7	Alcohol/daily		
8	Respirations pre-op		
9	Frequency of PA care		
10	Pulse pre-op		
11	Pre-op oral fluids		
12	Urinalysis-lab		
13	Clotting-time current		
14	Skin appearance		
15	Naso-gastric tube		
16	Wound-swab report		
17	? passed urine		
18	Haemoglobin-current		
19	Drainage tubes		
20	Type/admission		
21	Fluid chart		
22	Current oral fluids		
23	Indwelling catheter		
24	B/P pre-op		
25	Current diet		

Card Question		Order	Answer
26	Respirations current		
27	Level reported anxiety/ distress		
28	Frequency of dressings		
29	Socio-economic status		
30	Platelets current		
31	Marital status		
32	Skin to touch		
33	? bowels opened		
34	Time in recovery		
35	Duration of anaesthesia		
36	Cigarettes smoked daily		
37	Appearance/wound		
38	Appearance/abdomen		
39	Housing conditions		
40	Temperature current		
41	ESR current		
42	B/P current		
43	Current antibiotic		
44	Wound closure technique		
45	Vomiting		
46	Past medical history		
47	Expectoration		
48	Pre-med given		
49	Pre-surgical mobility		
50	Pulse current		

Card	Question	Order	Answer
51	Bowel sounds		
52	IVI		
53	Appearance of drainage		
54	Type of anaesthesia		
55	Weight		
56	Occupation		
57	Anti-emetic received		
58	Analgesia received		
59	Type/surgery		
60	Time since redressing		
61	Urinalysis ward		
62	Facial expression		
63	Pain now		
64	Cyanosis		
65	Gender		
66	Past psychiatric hist.		
67	Pre-op diet		
68	Time since surgery		
69	Muscle tone		
70	Quantity/drainage		
71	Current mobility		
72	Height		
73	Duration/surgery		
74	Degree/depression		
75	Temperature pre-op		
76	Nausea		

Diagnosis ______________________

APPENDIX 25

A. PARALYTIC ILEUS

GENERAL

1.	Gender	Male
2.	Age	24 years
3.	Weight	11 st
4.	Height	5' 10"
5.	Frame size	Medium
6.	Marital status	Single
7.	Socio-economic status	Single
8.	Housing conditions	Own flat
9.	Occupation	Accountant
10.	Past medical history	No previous hospital admission
11.	Past psychiatric history	None
12.	Cigarettes smoked daily	Does not smoke
13.	Alcohol consumed daily	1-2 pints beer
14.	Pre-surgical mobility	Full
15.	Type of admission	Emergency
16.	Disease state	Acute appendicitis
17.	Type of surgery	Appendectomy
18.	Pre-med given	Atropine lung Pethidine 100mgs
19.	Type of anaesthesia	General-Intubated-Muscle relaxed
20.	Duration of anaesthesia	1 hour
21.	Duration of surgery	1 hour
22.	Time in recovery	45 minutes
23.	Time since surgery	3 days
24.	Wound closure technique	Clips
25.	Frequency of dressings	Wound sprayed-Inspected daily
26.	Frequency of PA care	3-4 hly prn
27.	Drainage tubes	Nil
28.	IVI	Dextrose/saline + KVO
29.	Nasogastric tube	None
30.	Indwelling catheter	Nil
31.	Fluid chart	On fluid chart

SPECIFIC

32.	Level of consciousness	Alert
33.	Facial expression	Moderate distress
34.	Skin appearance	Pallor
35.	Skin to touch	Dry
36.	Muscle tone	Rigid abdomen
37.	Current mobility	Walks short distances
38.	Level of reported anxiety/distress	Moderate distress
39.	Degree of depression	None
40.	B/P pre-op	120/80
41.	B/P current	120/80
42.	Pulse pre-op	80
43.	Pulse current	90
44.	Temp. pre-op	Normal
45.	Temp. current	Normal
46.	Resp. pre-op	24
47.	Resp. current	30
48.	Cyanosis	None
49.	Expectoration	None
50.	Pre-op oral fluids	Nil wanted by patient
51.	Current oral fluids	30mls hourly
52.	Pre-op diet	Nil immediately before
53.	Current diet	H20 only
54.	Nausea	Recurrent. Relieved by vomiting
55.	Vomiting	x 6 since surgery
56.	Anti-emetic received	Maxalon x 2. Nil effect
57.	Bowel sounds	Nil
58.	? Bowels opened	No
59.	? Passed urine	400mls since surgery
60.	State of bladder	Not palpable or uncomfortable
61.	Pain now	Wound discomfort
62.	Analgesia received	Pethidine 100mgs x 2 since surgery
63.	Appearance of abdomen	Distension
64.	Appearance of wound	Satisfactory
65.	Appearance of dressing	Nil
66.	Time since re-dressed	Not dressed
67.	Appearance of drainage	Nil

68.	Quantity of drainage	Nil
69.	Current antibiotic	Ampicillin 250mgs hly, Cloxacillin 250mgs
70.	Urinalysis ward	+ve protein-nil else
71.	Urinalysis lab	No pathogens
72.	Wound swab report	Not taken
73.	Haemoglobin current	15g per 100ml
74.	ESR current	15mm in one hour
75.	Clotting time current	9.5 mins
76.	Platelets current	2.75,000 per c.mm

B. PULMONARY EMBOLISM

GENERAL

1.	Gender	Female
2.	Age	40 years
3.	Weight	11 stone
4.	Height	5' 3"
5.	Frame size	Medium
6.	Marital status	Married
7.	Socio-economic status	Good
8.	Housing conditions	Own home
9.	Occupation	Journalist
10.	Past medical history	1972-D & C, Recurrent bronchitis
11.	Past Psych. history	None
12.	Cigarettes smoked daily	40+
13.	Alcohol consumed daily	2-3 gins every day
14.	Pre-surgical mobility	Full
15.	Type of admission	Emergency
16.	Disease state	Acute appendicitis
17.	Type of surgery	Appendectomy
18.	Pre-med given	Omnipon 20mgs. Seopolamine 6mgs
19.	Type of anaesthesia	General, intubated
20.	Duration of anaesthesia	1 hour
21.	Duration of surgery	1 hour
22.	Time in recovery	1½ hours
23.	Time since surgery	24 hours
24.	Wound closure technique	Clips
25.	Frequency of dressings	Wound sprayed, daily inspection
26.	Frequency of PA care	3-4 hourly
27.	Drainage tubes	Nil
28.	IVI	Taken down 12 hours post-op
29.	Nasogastric tube	None
30.	Indwelling catheter	Nil
31.	Fluid chart	On fluid chart

SPECIFIC

1.	Level of consciousness	Feels faint
2.	Facial expression	Distress
3.	Skin appearance	Pallor
4.	Skin to touch	Cold and clammy
5.	Muscle tone	Normal
6.	Current mobility	Bed bound
7.	Level of reported anxiety/distress	Distressed
8.	Degree of depression	None
9.	B/P pre-op	140/80
10.	B/P current	100/60
11.	Pulse pre-op	78
12.	Pulse current	94
13.	Temp. pre-op	Normal
14.	Temp. current	35°c
15.	Resps. pre-op	24-28
16.	Resps current	32 Dyspnoea
17.	Cyanosis	Present
18.	Expectoration	Frothy haemoptysis
19.	Pre-op oral fluids	Free
20.	Current oral fluids	Free fluids
21.	Pre-op diet	Normal
22.	Current diet	Not resumed
23.	Nausea	None
24.	Vomiting	None
25.	Anti-emetic received	None
26.	Bowel sounds	Present
27.	? bowels opened	Not yet
28.	Passed urine	Has passed 400mls total
29.	State of bladder	NAD
30.	Pain now	Unilateral chest pain
31.	Analgesia received	100mgs Pethidine 1 hr ago, good effe
32.	Appearance of abdomen	Normal
33.	Appearance of wound	Satisfactory
34.	Appearance of dressing	No dressing
35.	Time since re-dressed	-

36. Appearance of drainage	No drainage
37. Quantity of drainage	Nil
38. Current antibiotic	Ampicillin 250mgs, Clox 250mgs 6 hly
39. Urinalysis-ward	Protein trace
40. Urinalysis-lab	No pathogens
41. Wound swab report	None taken
42. Haemoglobin-current	13g per 100ml
43. ESR - current	14mm in 1 hour
44. Clotting time - current	3.5 mins
45. Platelets - current	400,000 per c.mm

C. URINARY RETENTION

GENERAL

1.	Gender	Male
2.	Age	60 years
3.	Weight	12 stone
4.	Height	6'
5.	Frame sixe	Large
6.	Marital status	Married
7.	Socio-economic state	Semi-skilled
8.	Housing conditions	Local authority flat
9.	Occupation	Factory worker
10.	Past medical history	1976 myocardial infarct, no recurrence
11.	Past psych. history	NAD
12.	Cigarettes smoked daily	Does not smoke
13.	Alcohol consumed daily	Occasional beer
14.	Pre-surgical mobility	Full
15.	Type of admission	Emergency
16.	Disease state	Appendicitis
17.	Type of surgery	Appendectomy
18.	Pre-med given	20mg Omnipon, 1mg Atropine
19.	Type anaesthesia	General, intubated
20.	Duration anaesthesia	1 hour
21.	Duration surgery	1 hour
22.	Time in recovery	3 hours
23.	Time since surgery	16 hours
24.	Wound closure technique	Clips
25.	Frequency of dressings	Daily inspection
26.	Frequency of PA care	3-4 hourly
27.	Drainage tubes	None
28.	IVI	Dextrose/saline + KVO
29.	Nasogastirc tube	None
30.	Indwelling catheter	None
31.	Fluid chart	On fluid chart

SPECIFIC

1.	Level of consciousness	Alert
2.	Facial expression	Moderate distress
3.	Skin appearance	Normal
4.	Skin to touch	Slightly moist
5.	Muscle tone	Tense
6.	Current mobility	Up in chair for short periods
7.	Level of reported anxiety/distress	Some distress
8.	Degree of depression	None evident
9.	B/P pre-op	130/170
10.	B/P current	140/170
11.	Pulse pre-op	82
12.	Pulse current	90
13.	Temp. pre-op	35
14.	Temp. current	37
15.	Resps. pre-op	24
16.	Resps. current	30
17.	Cyanosis	None
18.	Expectoration	None
19.	Pre-op fluids	Free
20.	Current oral fluids	30mls per hour
21.	Pre-op diet	Normal
22.	Current diet	None
23.	Nausea	None
24.	Vomiting	None
25.	Anti-emetic received	None
26.	Bowel sounds	Present
27.	? bowels opened	Not yet
28.	? passed urine	Has not passed urine
29.	State of bladder	Bladder palpable
30.	Pain now	Abdominal pain
31.	Analgesia received	Pethidine 100mgs x 2 since surgery
32.	Appearance of abdomen	Distended
33.	Appearance of wound	NAD
34.	Appearance of dressing	Wound sprayed

35. Time since re-dressed	Daily inspection
36. Appearance of drainage	Nil
37. Quantity of drainage	Nil
38. Current antibiotic	Ampicillin 250mgs, Clox 250mgs 4 hly
39. Urinalysis - ward	+ve protein trace
40. Urinalysis - lab	No pathogens
41. Wound swab report	None sent
42. Haemoglobin - current	15g per 100ml
43. ESR - current	16.5mm in 1 hour
44. Clotting time - current	6 minutes
45. Platelets - current	320,000 per c. mm

D. WOUND INFECTION

GENERAL

1.	Gender	Male
2.	Age	74 years
3.	Weight	14 stone
4.	Height	5' 11"
5.	Frame size	Large
6.	Marital status	Widower
7.	Socio-economic state	Pensioner
8.	Housing conditions	Local authority flat
9.	Occupation	Retired bus driver
10.	Past medical history	History of intermit. pain L calf when walking, starting at 20yds
11.	Past Psych. history	Nil
12.	Cigarettes smoked daily	30-40
13.	Alcohol consumed daily	1 pint
14.	Pre-surgical mobility	Walk 20yds
15.	Type of admission	Booked from W.L.
16.	Disease state	Intermittend claudication L
17.	Type of surgery	Lumbar sympathectomy L
18.	Pre-med given	Omnipon 20mgs, Scopolamine
19.	Type anaesthesia	General. Intubated
20.	Duration of anaesthesia	1 hour
21.	Duration of surgery	1 hour
22.	Time in recovery	2 hours
23.	Time since surgery	4 days
24.	Wound closure technique	Black silk sutures
25.	Frequency of dressings	x 2 daily
26.	Frequency of PA care	pnn
27.	Drainage tubes	Nil
28.	IVI	Nil
29.	Nasogastric tube	Nil
30.	Indwelling catheter	Nil
31.	Fluid chart	Nil

SPECIFIC

1.	Level of consciousness	Alert
2.	Facial expression	Does not indicate distress
3.	Skin appearance	NAD. Left foot pink
4.	Skin to touch	NAD. Left foot warm
5.	Muscle tone	NAD
6.	Current mobility	Walking in the ward to loo & dayroom
7.	Level of reported anxiety/distress	Slight anxiety
8.	Degree of depression	Nil
9.	B/P pre-op	160/90
10.	B/P current	160/90
11.	Pulse pre-op	76
12.	Pulse current	76-80
13.	Temp. pre-op	TN
14.	Temp. current	37.5
15.	Resps. pre-op	24
16.	Resps. current	24
17.	Cyanosis	None
18.	Expectoration	Mucoid
19.	Pre-op oral fluids	Free fluids
20.	Current oral fluids	Free fluids
21.	Pre-op diet	Normal
22.	Current diet	Normal
23.	Nausea	None
24.	Vomiting	None
25.	Anti-emetic received	None
26.	Bowel sounds	Present
27.	? Bowels opened	Yes
28.	? passed urine	passed urine normally
29.	State of bladder	Not palpable
30.	Pain now	Some wound discomfort
31.	Analgesia received	Paracetamol prn/4 hlr
32.	Appearance of abdomen	Not distended
33.	Appearance of wound	Inflammed edges

34. Appearance of dressing	Mod. discharge of purulent
35. Time since re-dressed	8 hours
36. Appearance of drainage	Nil
37. Quantity of drainage	Nil
38. Current antibiotic	Ampicillin 250mgs, Clox. 250 oral
39. Urinalysis - ward	NAD
40. Urinalysis - lab	Not performed
41. Wound swab report	Swab sent to lab this am
42. Haemoglobin - current	15g per 100ml
43. ESR - current	16mm in 1 hour
44. Clotting time - current	7 minutes
45. Platelets - current	250,000 per c. mm

APPENDIX 26

TASK INSTRUCTIONS

In front of you are two packs of cards and two lists, please do not touch them at present. I want you to imagine that you are caring for a general surgical patient post-operatively. Your task is to diagnose/judge the current condition of this patient. I want you to do this twice and on each occasion the patient's condition will be different. I will now give you a list of the possible patient conditions that you could diagnose. The diagnosis that you make will be one of the conditions on that list. When you have finished reading the patient conditions please take the list that is on top of the materials in front of you.

At this point I want to ask you NOT to look at what your neighbours are doing. At no point will two people sitting next to each other be doing the same task. That is, the diagnoses that you will make will be different. I should also like to ask you not to talk at any point until we have finished. The reason for this is that different people have different ways of doing things and what is useful for one person may confuse someone else. What is more useful for you and for me is to do these tasks as you would do them naturally, other peoples' strategies may cause you to make a mistake.

As you can see your lists contain a series of questions about information that you may wishe to know about your patient. There is no logic in the order of the questions so do not start with question 1 unless you wish to know the answer to that question, start anywhere in the list. So decide what you first want to know about your patient on your list and put a number 1 in the column marked order beside that question. Now take the top pack of cards in front of you, the question that you wish to ask has a list number beside it. Take the card with that list number on it from your pack - turn it over and you have the answer. write that answer down on your list beside the question that you asked. Do not look at any other card other than the question that you are currently asking.

Having recorded the answer to question 1 then proceed to number 2 and put a 2 in the column beside that question - look at the appropriate card and record the answer. It is important for me to know the order in which you ask each question, so please do not forget to write 2 for the second, 3 for the third etc. Then ask a third question and look at each card in turn, as you ask your question. Please do not look at any card that you are not specifically interested in. On one occasion you can ask as many questions as you like, on the other you are limited to 12 questions. When you are limited to 12 questions this is written on the top of the list. Please do exactly that and just ask 12 questions when instructed to do so and as many as you like when not so instructed. Not all the questions will be relevant, do not ask any more than you feel you need to. St start now by asking the questions, one by one.

Record the answer and the order used.

When you have finished a task write your diagnosis for your patient in the space provided at the bottom of your list. When you have done this do your 2nd task, either asking 12 questions or as many as you want to according to any instructions at the top of the sheet. The two lists are different - please do not worry about this. These first two tasks are for you to practice so that you get the hang of what to do, when you have finished I will give you two more task packs. Do exactly the same again.

If you have any worries ask me to come over and whisper your query to me.

Please do this all in silence - it doesn't in fact take all that long and do not look at your neighbours either side.

I'll tell you exactly what I'm up to and answer any questions when you have finished.

APPENDIX 27

Anova Summary Table - Task I

Source of Variance	SS	DF	MS	F	P.
Main effects	138.943	3	46.314	3.896	0.012
Task (cord)	63.920	1	63.920	5.377	0.023
Pool (small or large)	4.102	1	4.102	0.345	0.559
Restriction	70.920	1	70.920	5.965	0.017
2-way interaction	48.580	3	16.193	1.362	0.260
Task x pool	0.557	1	0.557	0.047	0.829
Task x restriction	48.011	1	48.011	4.038	0.048
Pool x restriction	0.011	1	0.011	0.001	0.975
3-way interaction	0.102	1	0.102	0.009	0.926
Task x pool x restr.	0.102	1	0.102	0.009	0.926
Explained	187.625	7	26.804	2.255	0.038
Residual	951.091	80	11.889		
Total	1138.716	87	13.089		

Anova Summary Table - Task II

Source of Variance	SS	DF	MS	F	P.
Main effects	245.125	3	81.708	5.745	0.001
Task (Cord)	8.284	1	8.284	0.582	0.448
Pool (small or large)	17.284	1	17.284	1.215	0.274
Restriction	219.557	1	219.557	15.437	0.000
2-way interactions	6.125	3	2.042	0.144	0.934
Task x pool	0.920	1	0.920	0.065	0.800
Task x restriction	3.284	1	3.284	0.231	0.632
Pool x restriction	1.920	1	1.920	0.135	0.714
3-way interaction	1.375	1	1.375	0.097	0.757
Task x pool x restr.	1.375	1	1.375	0.097	0.757
Explained	252.625	7	36.089	2.537	0.021
Residual	1137.818	80	14.223		
Total	1390.443	87	15.982		

APPENDIX 28

Unrestricted Condition Both Pools

Frequency Distribution of Total Question use by Subjects as a Function of Stage in the Task when General Questions were Asked

Mean = 12.1
Median = 11
S.D. = 4.46

Mean = 15.56
Median = 15
S.D. = 4.66

APPENDIX 29

Unrestricted/Both Pools

Anova Least Squares Solution. Total Number of Questions Asked by Subjects Asking General Questions Earlier of Subject Asking Later

	Right b_1							Wrong b_2	
LATE A_1	23	5	14	22				16	11
	15	15	21	16				24	23
	17	9	19	19				14	13
	12	12	12					15	23
	12	11	11					18	
	17	15	13						
EARLY A_2	21	11	14	21	10	7	10		
	7	9	23	9	6	17	8	11	9
	14	11	5	11	9	14		14	9
	17	9	13	14	5	10		12	19
	11	18	18	10	12	14		18	12
	8	9	19	18	9	10		20	11
	12	7	8	11	6	13			

Cell Frequencies

	b_1	b_2	
a_1	21	9	$n_1=30$
a_2	44	10	$n_2=54$
	$n_1=65$	$n_2=19$	84

Cell Totals

a_1	$AB_1=310$	$AB_2=157$	$A_1=467$
a_2	$AB_3=519$	$AB_4=135$	$A_2=654$
	$B_1=829$	$B_2=292$	$G=1121$

(1) = G^2/n = 14960

(2) = ΣX^2 = 16877 (S.D, 4.8)

(3) = $\Sigma(Ak^2/ni)$ = 15190.3

(4) = $\Sigma(Bj^2/nj)$ = 15060.5

(5) = $\Sigma\Sigma(ABij^2/nij)$ = 15259.3

Cell Means

	b_1	b_2		
a_1	14.8	17.4		
a_2	11.8	13.5		
dj	3	3.9		
wj	14.2	4.7	Σwj =	18.9
wjdj	42.6	18.33	Σwjdj =	60.93

$$SSab(adj) = \Sigma wjdj^2 - \frac{(\Sigma wjdj)^2}{wj}$$

$$\Sigma wjdj^2 = (3)(42.6) = (3.9)(18.33) = 199.29$$

$$SSab(adj) = 199.29 - \frac{(60.93)^2}{18.9} = 2.86$$

SS cells = (5) - (1) = 299.3
SSa = (3) - (1) = 230.3
SSb = (4) - (1) = 100.5

SSa(adj) = SS cells - SSab(adj) - SSb
= 299.3 - 2.86 - 100.5 = 195.94

SSb(adj) = SS cells - SSab(adj) - SSa - 299.3 - 2.86 - 230.3
= 66.14

SS error = (2)-(5) = 1617.7

Source of variation	SS	df	MS	VR	p<
Stage A	195.94	1	195.94	9.69	df 1.60 $p < 0.01$
Accuracy B	66.14	1	66.14	3.3	0.10
AB	2.86	2	1.43	.070	n.s
Error	1617.7	80	20.22		

The observed value of A, 9.69, exceeds the table value of F = 7.08 df 1.60 ∴ the Null hypothesis can be rejected at p 0.01 level of α .

The observed value of B, 3.3, exceeds the table value of F = 2.79 df 1.60 p 0.10 does not allow rejection of the Null hypothesis but is indicative of a trend at the 10% level.

APPENDIX 30

Unrestricted - Both Pools

X^2 Comparing the number of subjects asking questions - general, context questions in positions 1-3 (early) more often than other positions of those asking them more often later as a function of accuracy.

	Early	Later	
Correct	44 41.8	21 23.2	65
Incorrect	10 12.2	9 6.8	19
	54	30	84

$$X^2 = \frac{N(1AD-BC1-\frac{N}{2})^2}{(A+B)(C+D)(A+C)(B+D)}$$

$X^2 = .87$ dfl

From tables $X^2 > 2.71$ where $p < 0.10$ df 1, $\therefore$ we cannot reject the Null hypothesis with observed value of .87.

Appendix 31

Large Pool Restricted to 12 Questions: Frequency Distribution of Percentage of General Question use in Both Halves of Task I & II for all Subjects.

APPENDIX 32

General Context Question Use in the Second Half of Both Tasks - Restricted Large Pool, t Statistic

Incorrect		Correct	
9.09	0	0	0
25.0	0	0	0
8.3	14.3	0	16.7
0	0	0	8.3
0	8.3	0	0
0	0	8.3	
8.3	0	0	
8.3	0	0	
8.3	0	0	
9.09	16.7	0	
8.3	8.3	0	
0	0	0	
8.3	9.09	0	

$N_1 = 13$ $\quad N_2 = 31$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{\Sigma X_1^2 - \dfrac{(\Sigma X_1)^2}{n1} + X_2^2 - \dfrac{(\Sigma X_2)^2}{n1}}{(n_1-1) + (n_2-1)} \times \dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

$t = 2.23$ df $(n_1-1) + (n_2-1)$ 42

From table

α .025 the value of $t > 2.021$ df 40

2α .05

$\therefore$ the observed value $t = 2.23$ can reject the Ho at $p < 0.025$ for a one sided test.

APPENDIX 33

Restricted Large Pool. General Context Question use in the First Half of Task I and II.

Incorrect	Correct		
9.09	36.4	8.3	18.2
25	25	16.7	18.2
16.7	25	25	14.3
33.3	16.7	22.2	25
33.3	25	11.1	0
10		8.3	22.2
16.7		25	16.7
16.7		33.3	16.7
16.7		22.2	25
0		8.3	0
16.7		50	16.7
16.7		22.2	25
8.3		40	27.3

$N_1 = 13$ $N_2 = 31$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{\Sigma X_1^2 - \dfrac{(\Sigma X_1)^2}{n_1} + \Sigma X_1^2 - \dfrac{(\Sigma X_2)^2}{n_2}}{(n_1-1) + (n_2-1)} \times \dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

$t = 1.15$ df $(n_1-1) + (n_2-1) = 42$

Table value for one-sided test at the $p < 0.10$ trend level $t = 1.303$ df (40).

APPENDIX 34

Unrestricted Large Pool. General Context Question use in the Second Half of Both Tasks. t Statistics.

Incorrect	Correct				
0	4.3	5.9	0	8.3	11.1
0	0	0	20.0	9.09	0
25.0	6.7	8.3	0	11.1	7.1
20.8	11.8	25.0	22.2	0	8.7
14.3	0	5.9	0	14.3	20
0	0	0	6.7	0	5.3
8.3	0	0	11.1	4.8	7.7

n = 7 n = 36

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{\Sigma X_1^2 - \dfrac{(\Sigma X_1)^2}{n} + \Sigma X_2^2 - \dfrac{(\Sigma X_2)^2}{n}}{(n_1-1) \quad (n_2-1)} \times \dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

$t = 1.0476$ df $(n_1-1) + (n_2-1) = 41$

Table value α (one-side) $p < 0.05$, $t > 1.684$ df 40
$p < 0.25$, $t > 0.681$ df 40 & $p < .1$, $t > 1.303$

APPENDIX 35

Unrestricted Large Pool. General Context Question use in the First Half of Both Tasks. t Statistic

Incorrect	Correct					
9.09	8.7	11.8	0	25	22.2
0	9.5	18.2	20	0	28.6
12.5	6.7	41.7	18.2	44.4	14.3
12.5	0	25	33.3	13.3	13.04
21.4	42.8	35.3	18.2	7.1	20
14.3	0	37.5	13.3	11.1	15.8
16.7	21.4	16.7	11.1	14.3	15.4
					16.7
n = 7			n = 36		

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{\Sigma X_1^2 - \dfrac{(\Sigma X_1)^2}{n_1} + \Sigma X_2^2 - \dfrac{(\Sigma X_2)^2}{n_2}}{(n_1-1) + (n_2-1)} \times \dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

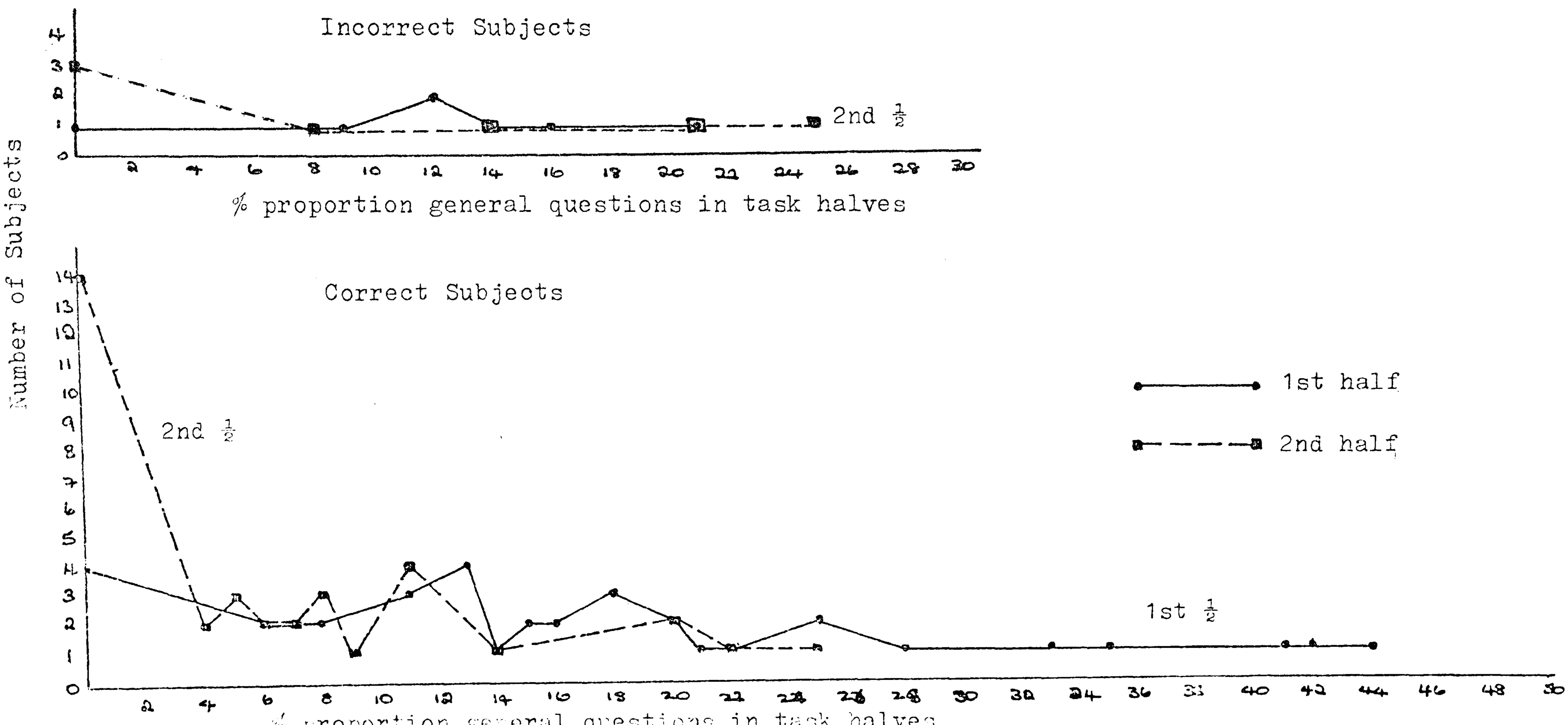
$t = 1.224$ df $(n_1-1) + (n_2-2) = 41$

From table $t > 1.303$ df 40, $p < 0.1$ for a one-sided test (α), $\therefore$ with an observed $t = 1.224$ the Ho is accepted at the 5% level, and no trend is observed.

APPENDIX 36

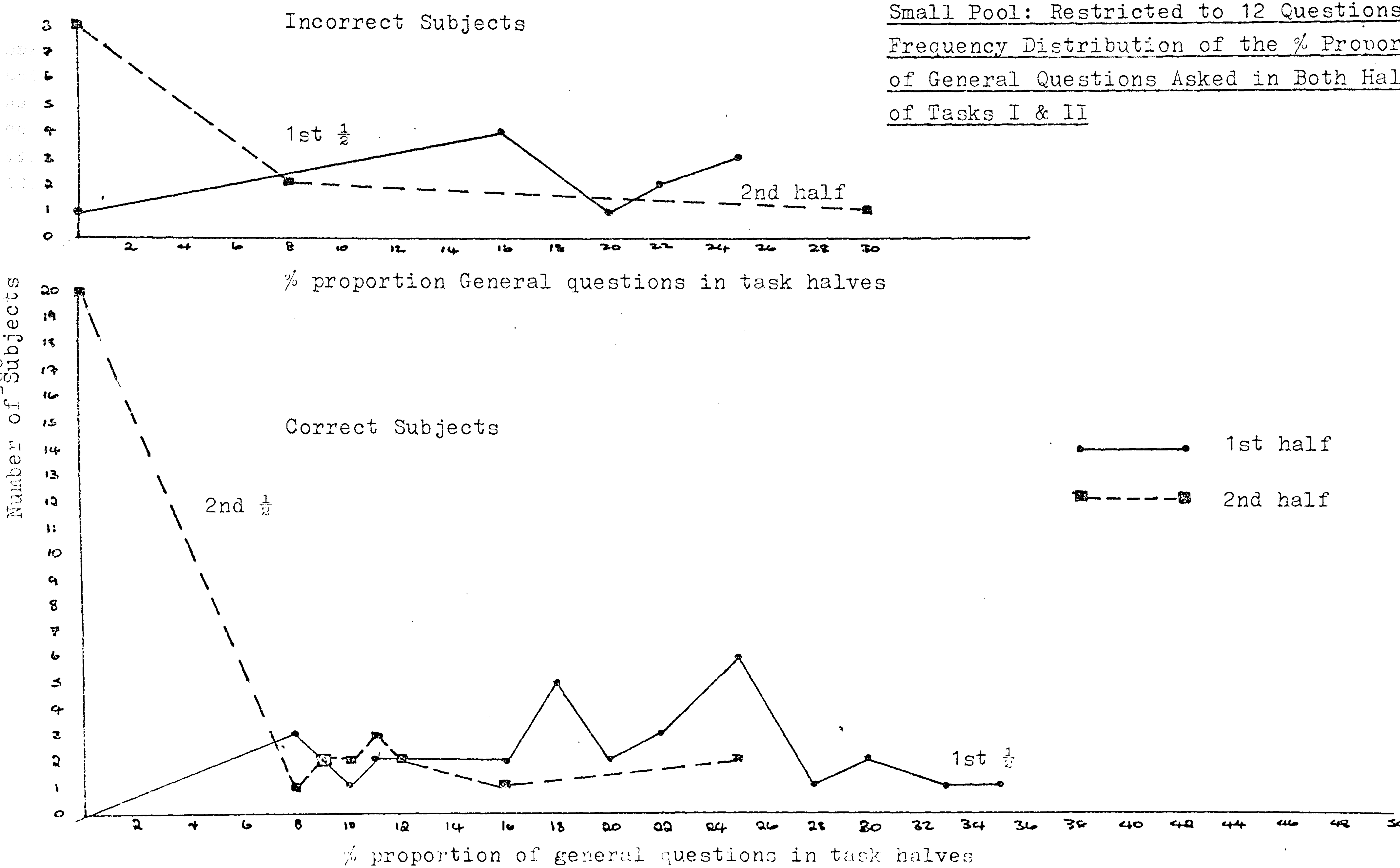
Unrestricted: Large Pool

Frequency Distribution of % Proportion of General Questions in Both Halves of Task I & II



APPENDIX 37

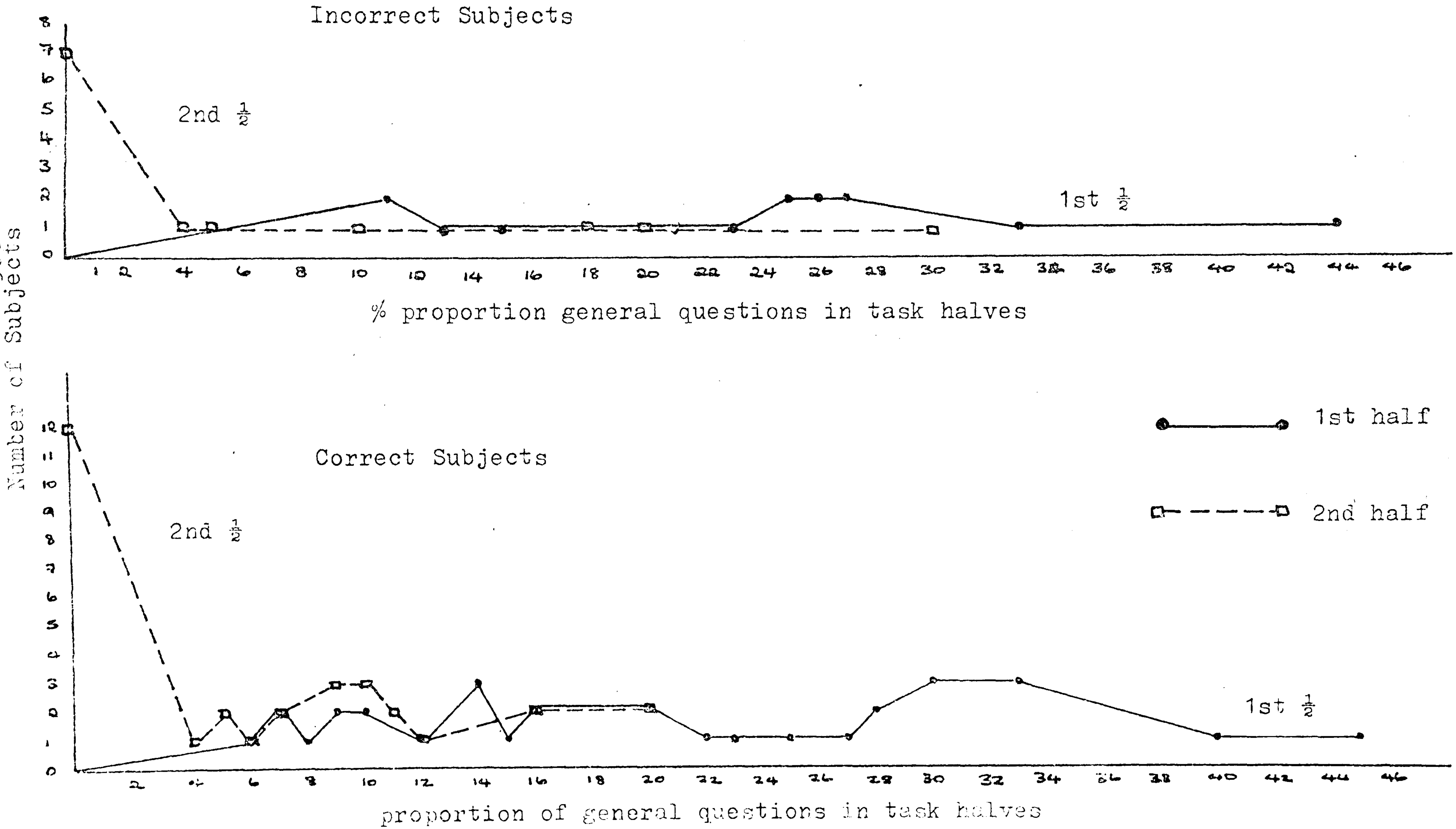
Small Pool: Restricted to 12 Questions Frequency Distribution of the % Proportion of General Questions Asked in Both Halves of Tasks I & II



Small Pool: Unrestricted

APPENDIX 38

Frequency Distribution of the % Proportion of General Questions Asked in Both Halves of Tasks I & II



APPENDIX 39

Name........................Ward..............Date............

Below are a number of words which describe moods. Please put a cross to indicate how much you have felt the way described today.

	Not at all	Little	Quite a bit	Extremely
Shaky				
Sluggish				
Resentful				
Nervous				
Weary				
Vigorous				
Hopeless				
Lively				
Guilty				
Tired				
Unhappy				
Tense				
Full of pep				
Active				
Worthless				
Miserable				
Worn out				
Discouraged				
Spiteful				
Depressed				
On edge				
Angry				
Furious				
Helpless				

APPENDIX 40

Name ....Ward..............Date..........

Below are a number of words which describe different moods. Please put a cross in the appropriate column beside the words to indicate how much you think each word describes how this patient felt today.

	Not at all	Little	Quite a bit	Extremely
Shaky				
Sluggish				
Resentful				
Nervous				
Weary				
Vigorous				
Hopeless				
Lively				
Guilty				
Tired				
Unhappy				
Tense				
Full of pep				
Active				
Worthless				
Miserable				
Worn out				
Discouraged				
Spiteful				
Depressed				
On edge				
Angry				
Furious				
Helpless				

APPENDIX 41

Name..........................Ward.............Date...........

Imagine yourself in this patient's position. Please rate this list of words as you would expect to feel if you were this patient.

	Not at all	Little	Quite a bit	Extremely
Shaky				
Sluggish				
Resentful				
Nervous				
Weary				
Vigorous				
Hopeless				
Lively				
Guilty				
Tired				
Unhappy				
Tense				
Full of pep				
Active				
Worthless				
Miserable				
Worn out				
Discouraged				
Spiteful				
Depressed				
On edge				
Angry				
Furious				
Helpless				

APPENDIX 42

Name..........................Ward...........'..Date..........

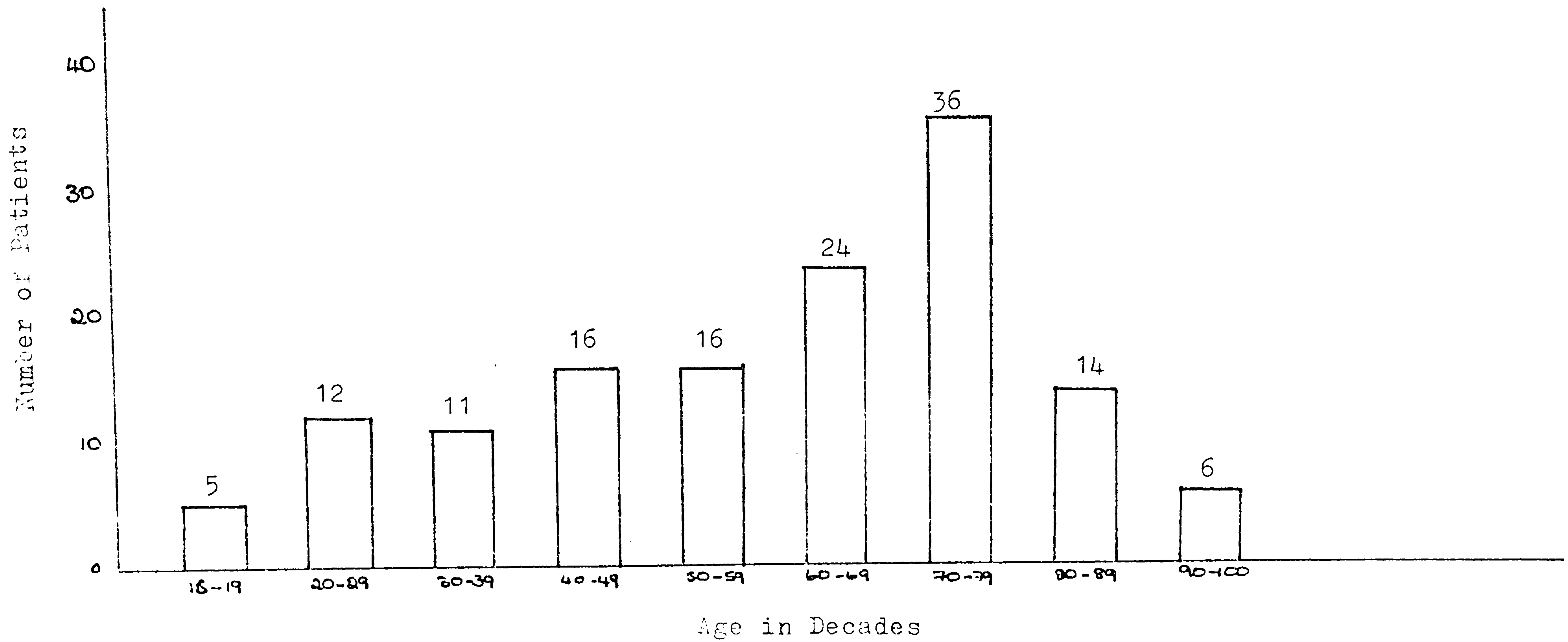
Imagine a typical patient of the same sex suffering from the same condition. Please rate this list of words as you would expect a typical patient to feel.

	Not at all	Little	Quite a bit	Extremely
Shaky				
Sluggish				
Resentful				
Nervous				
Weary				
Vigorous				
Hopeless				
Lively				
Guilty				
Tired				
Unhappy				
Tense				
Full of pep				
Active				
Worthless				
Miserable				
Worn out				
Discouraged				
Spiteful				
Depressed				
On edge				
Angry				
Furious				
Helpless				

APPENDIX 43

Age Distribution of the Patients

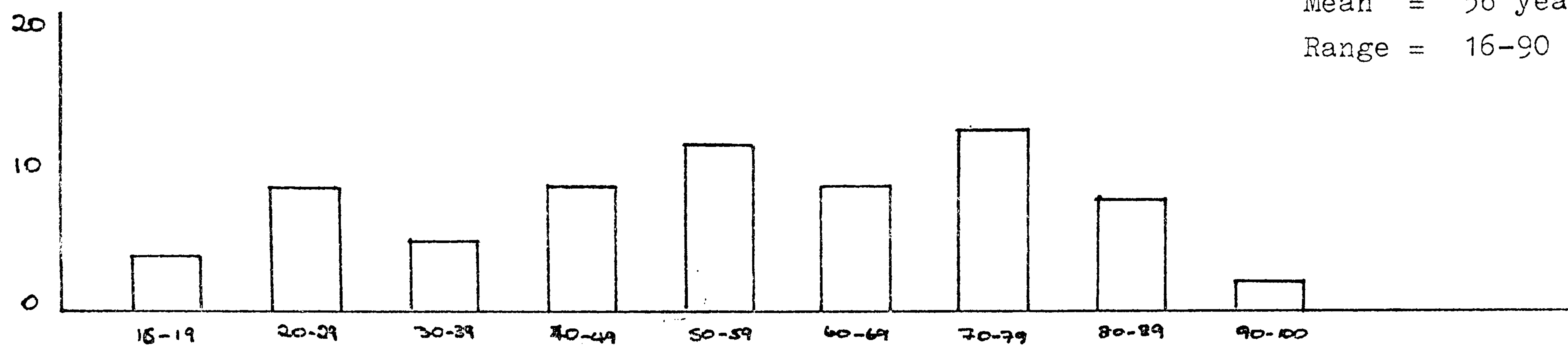
N = 140
Mean = 60 years
Range = 16-96 years



APPENDIX 44

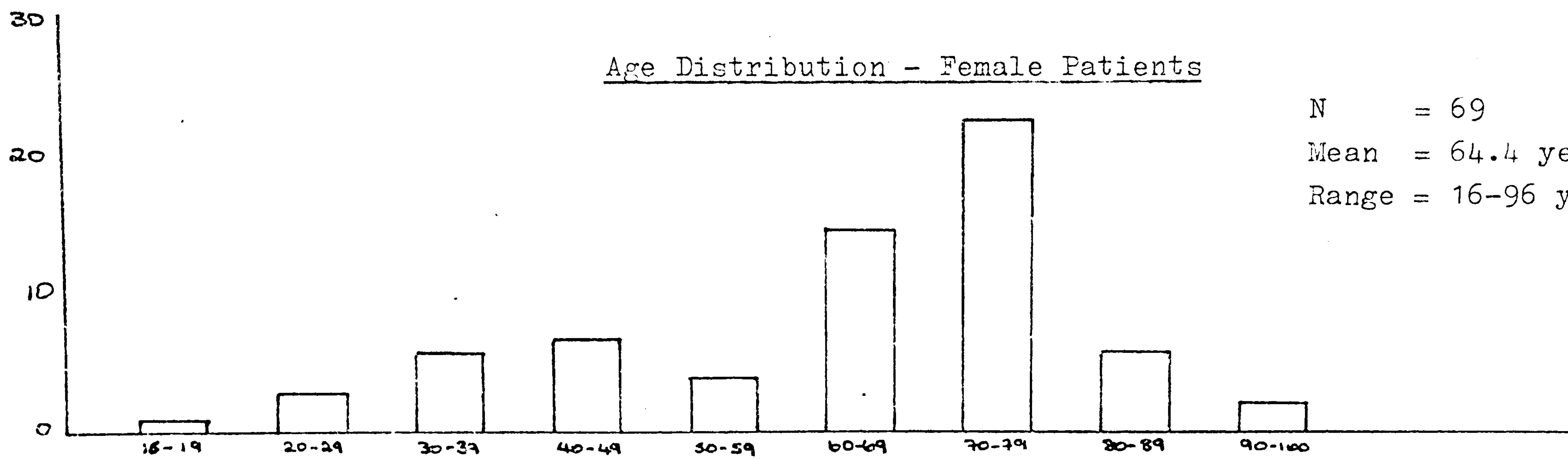
Age Distribution - Male Patients

N = 71
Mean = 56 years
Range = 16-90 years



Age Distribution - Female Patients

N = 69
Mean = 64.4 years
Range = 16-96 years



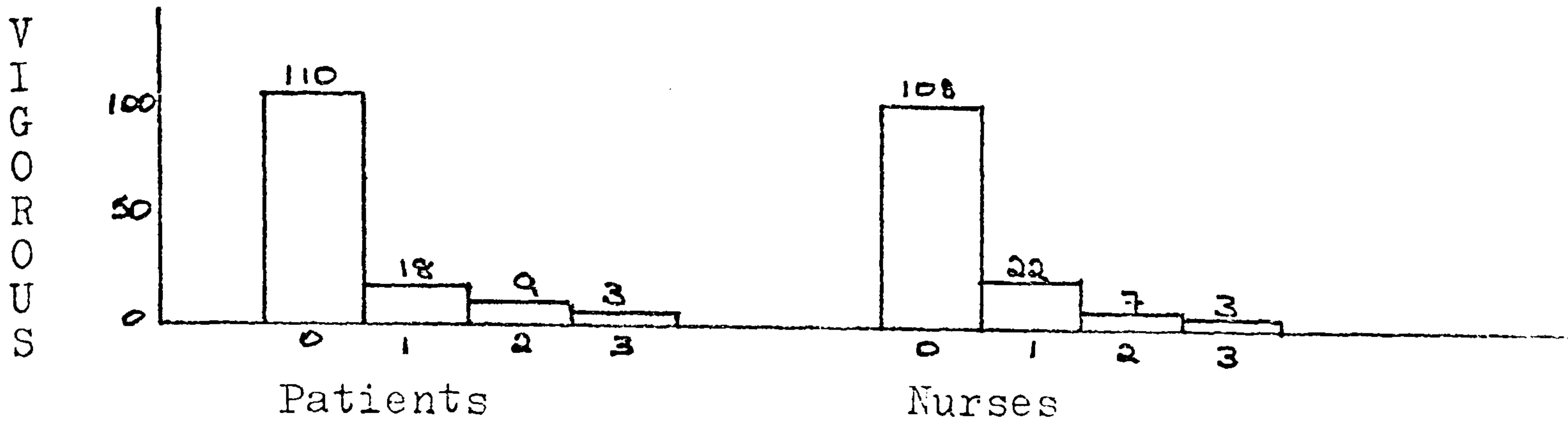
Factor I Vigor

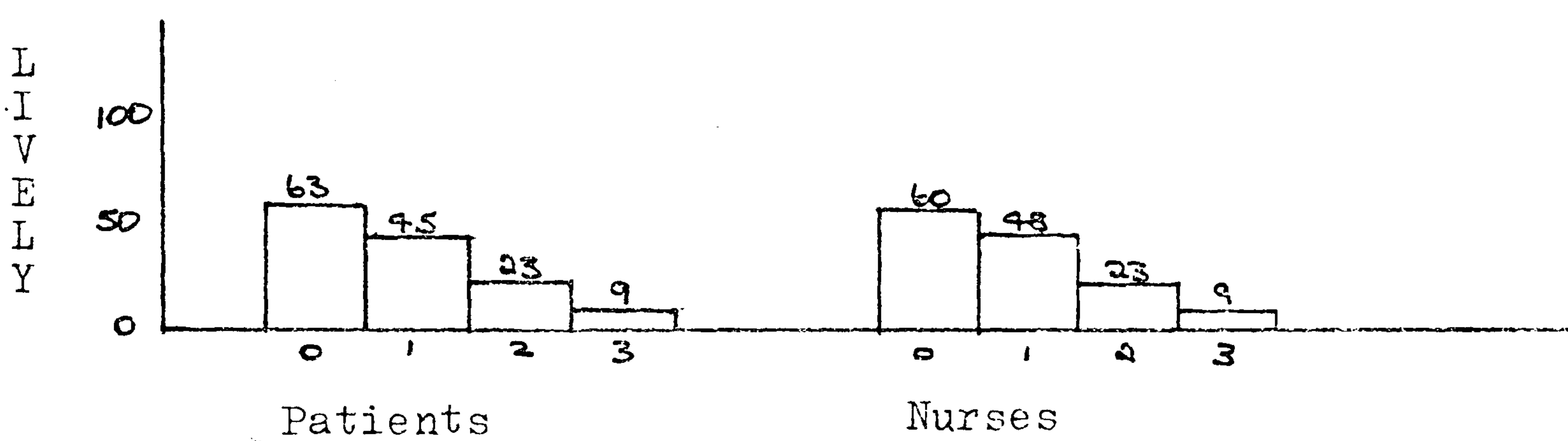
Cumulative Distribution of Nurses'/Patients Ratings

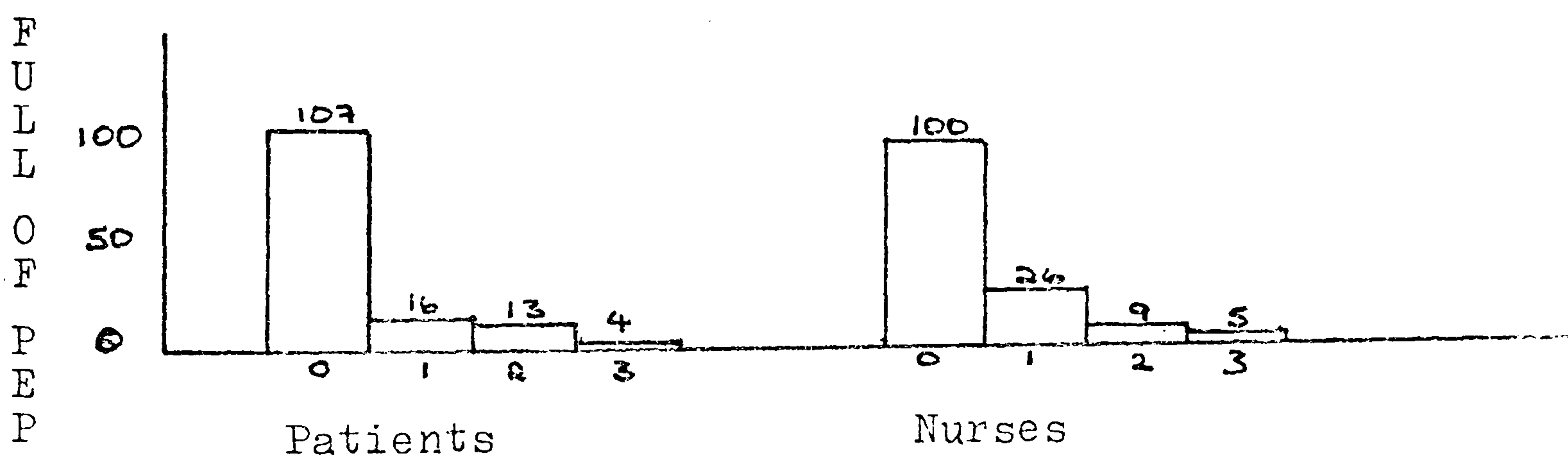
The Numbers of Patients and Nurses Rating Each Word

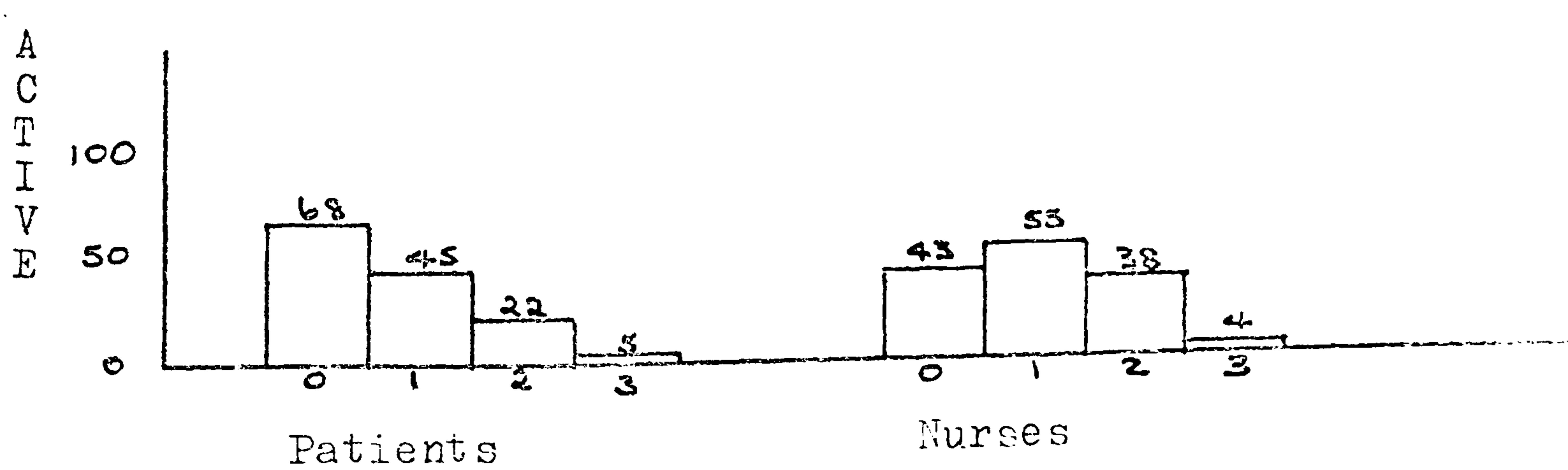
Nurses N = 140

Patients N = 140







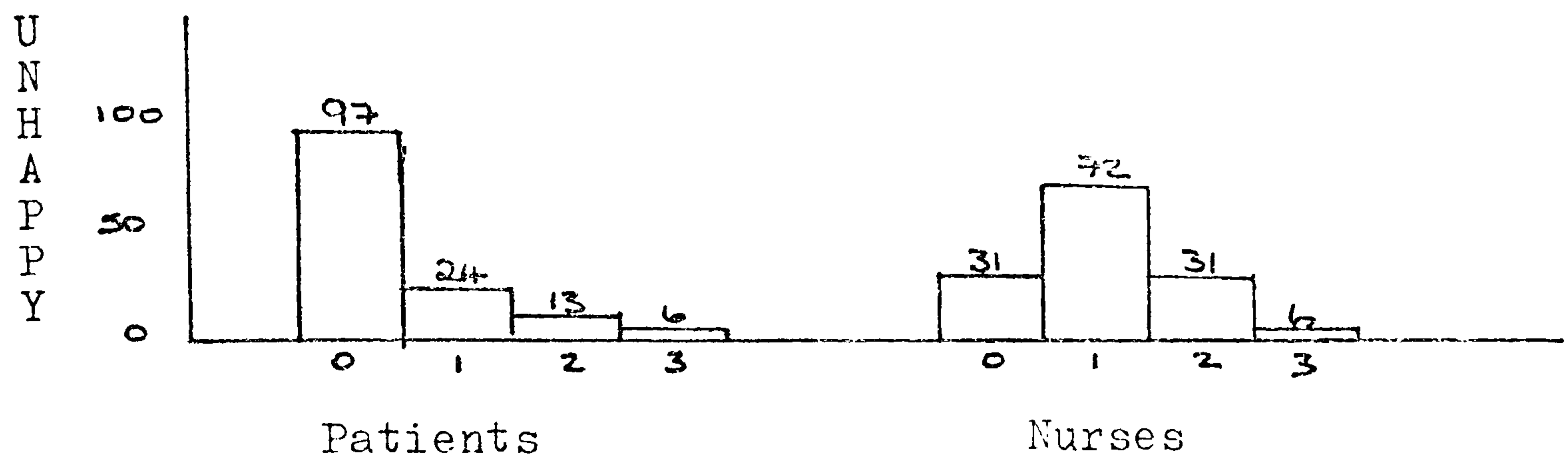


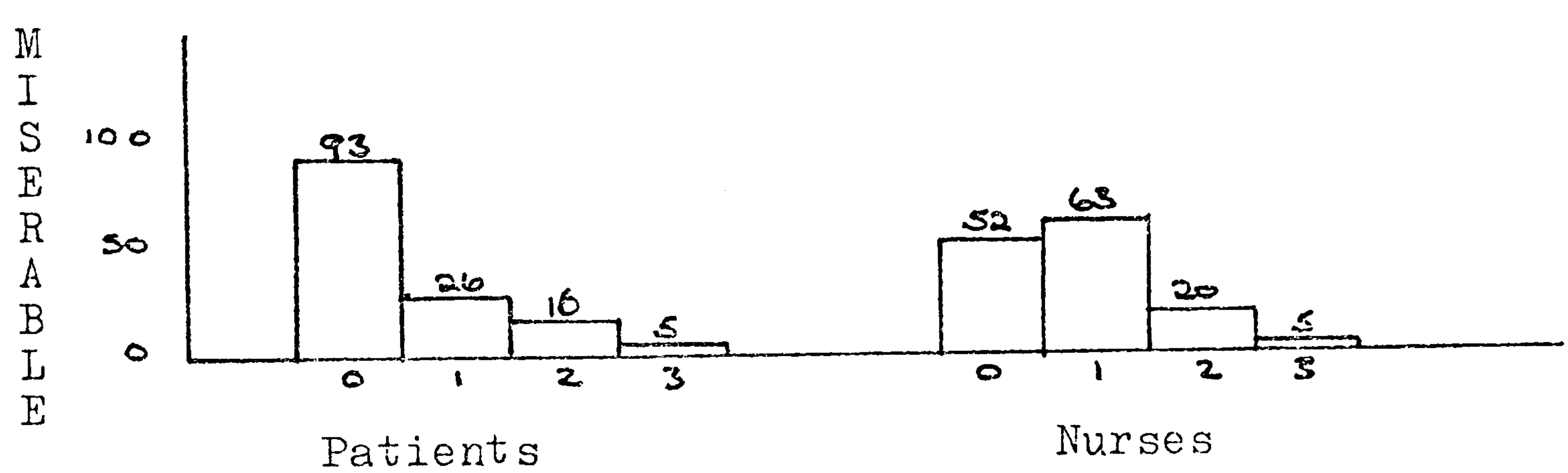
Cumulative Distribution of Nurses/Patients Ratings

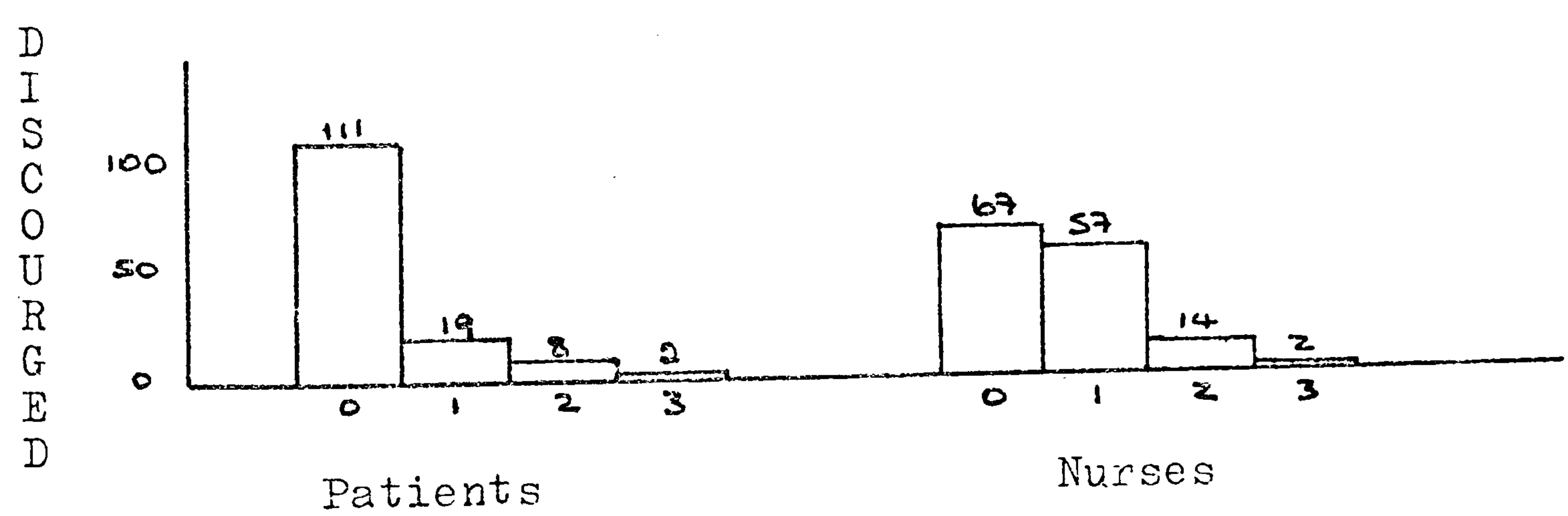
The Numbers of Patients and Nurses Rating Each Word

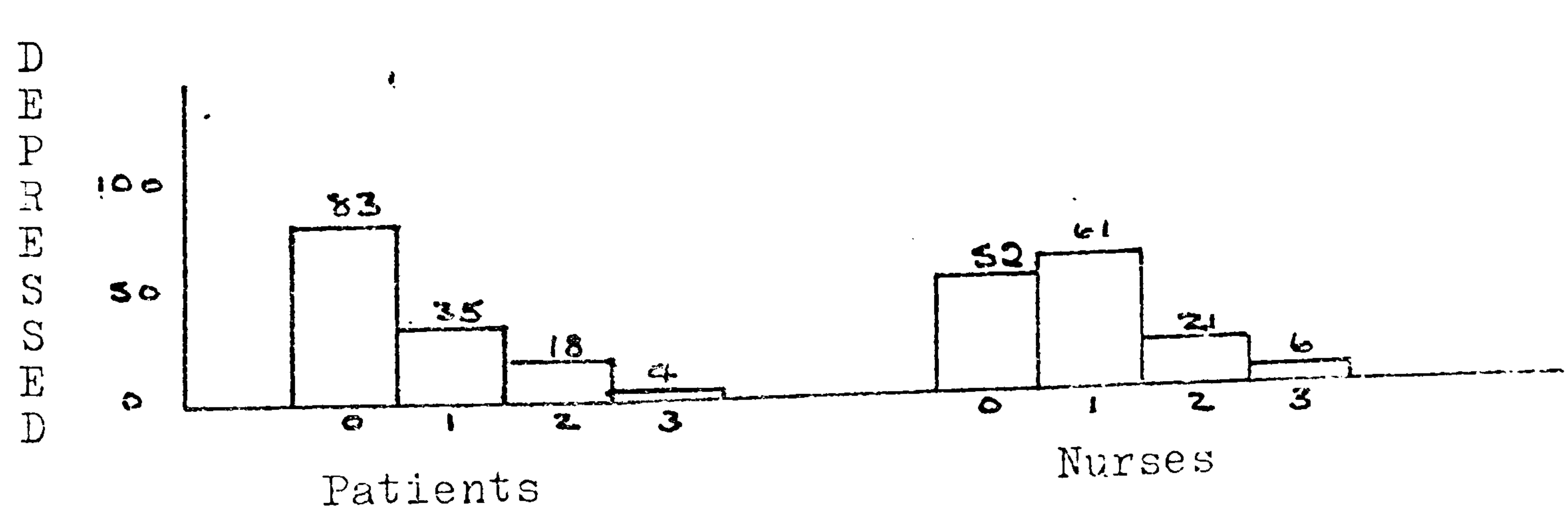
Nurses N = 140

Patients N = 140









Factor 3 Anxiety

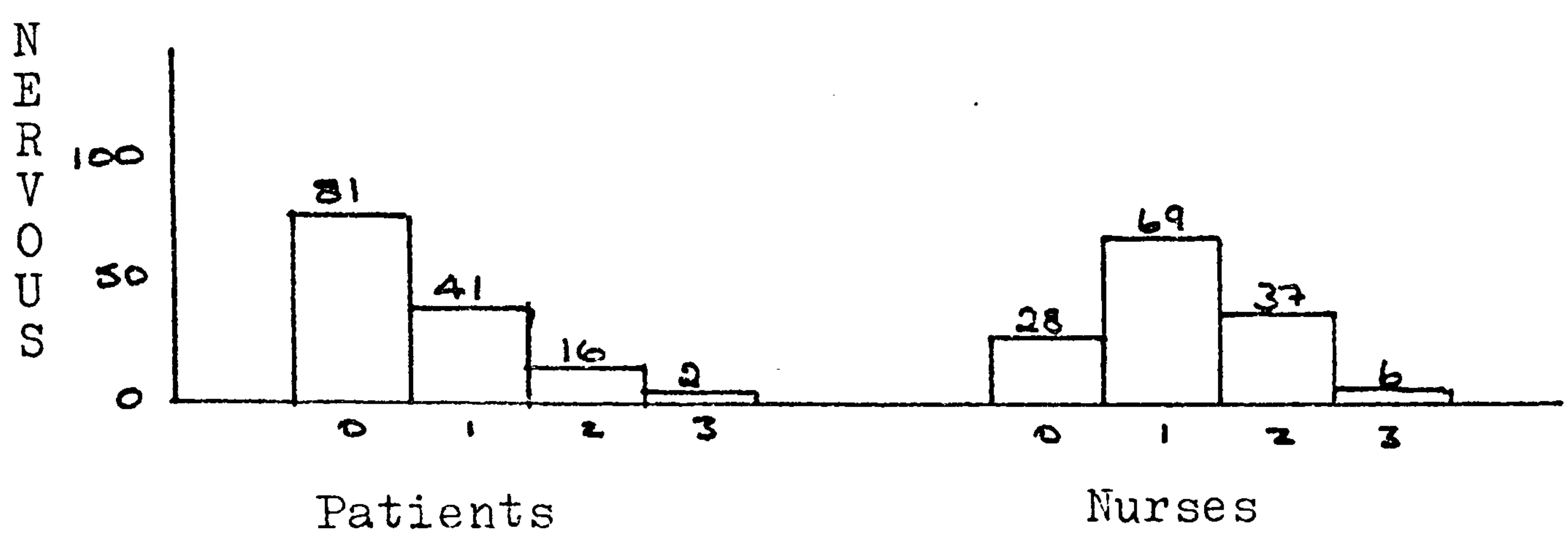
APPENDIX 47

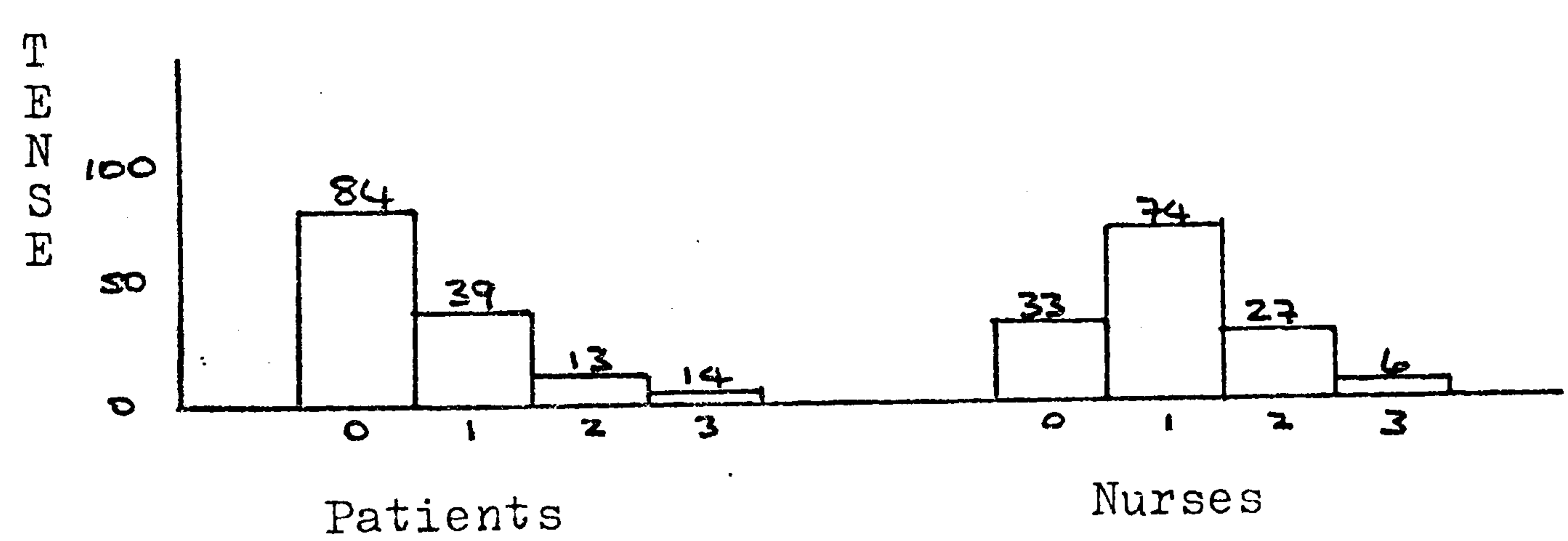
Cumulative Distribution of Nurses/Patients Ratings

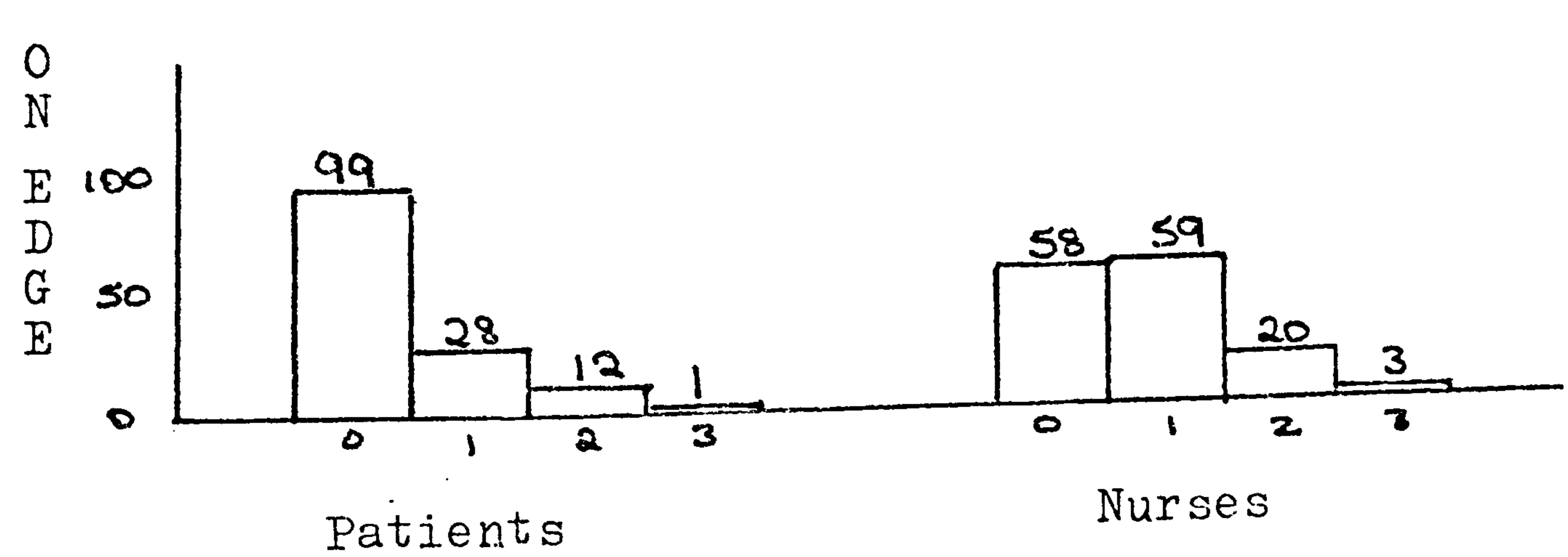
The Numbers of Patients and Nurses Rating Each Word

Nurses N = 140

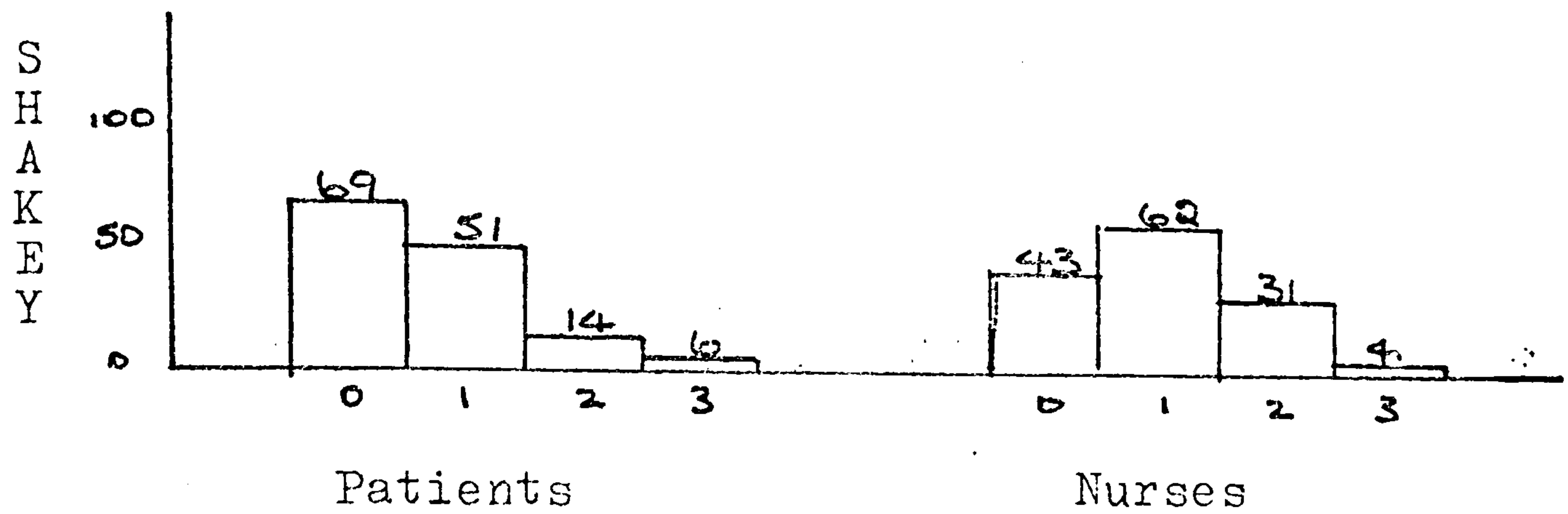
Patients N = 140

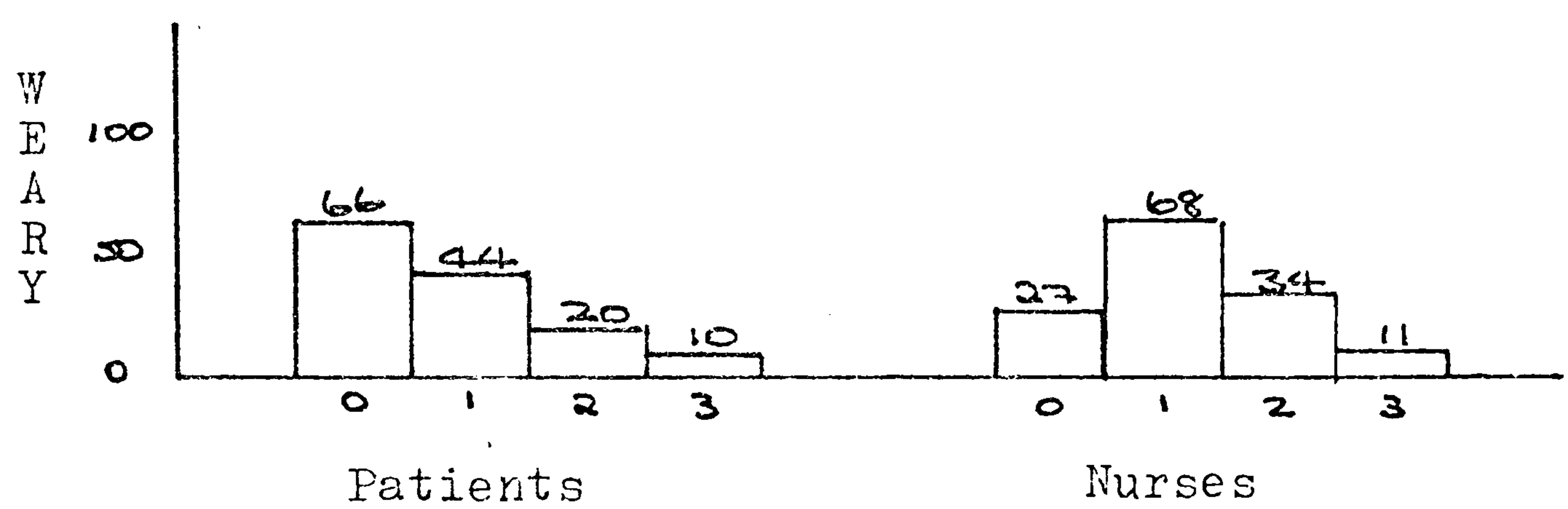


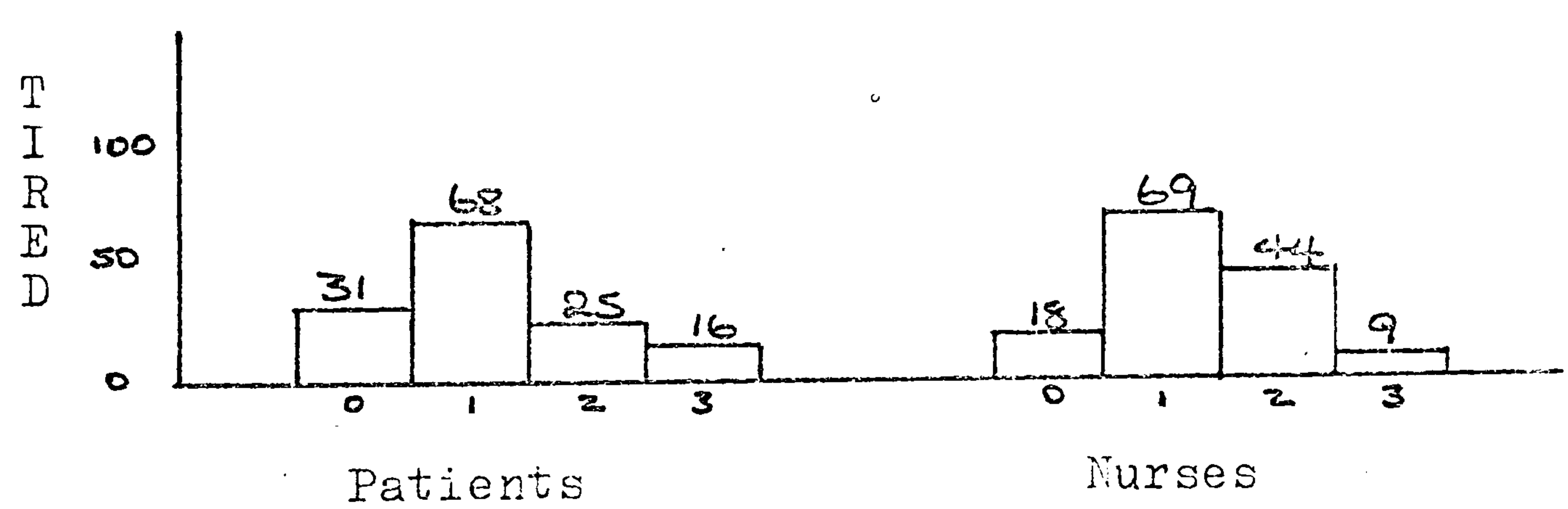


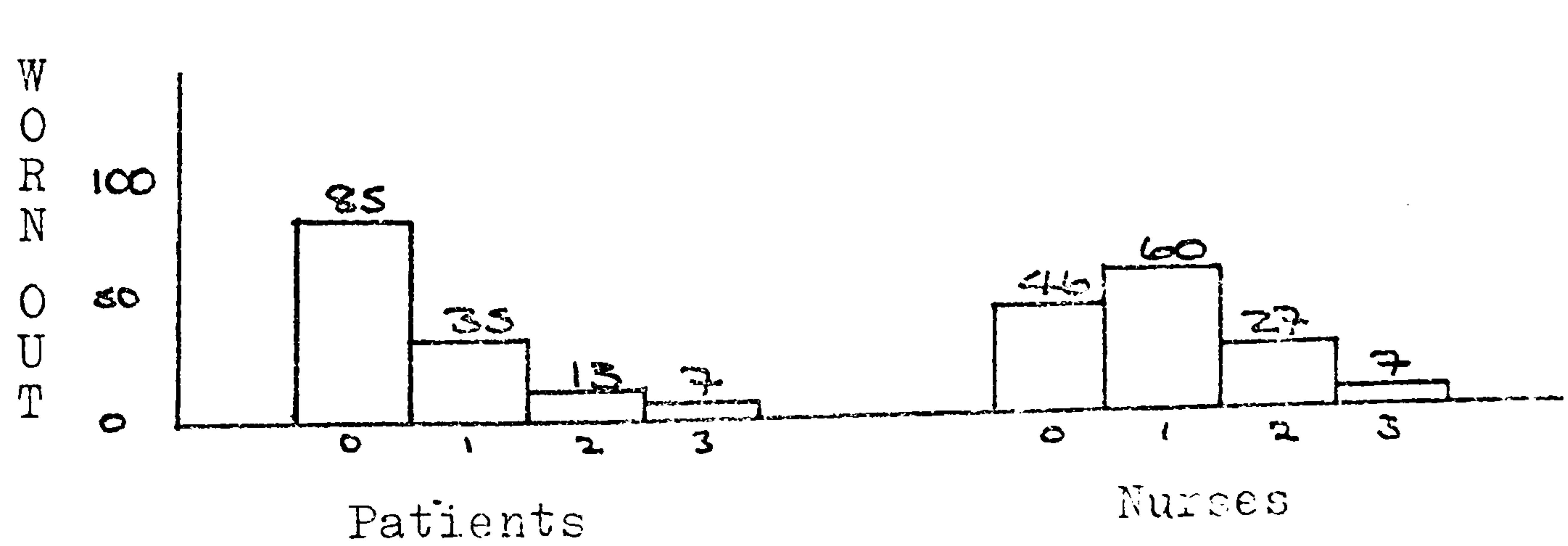


Factor 4 Fatigue

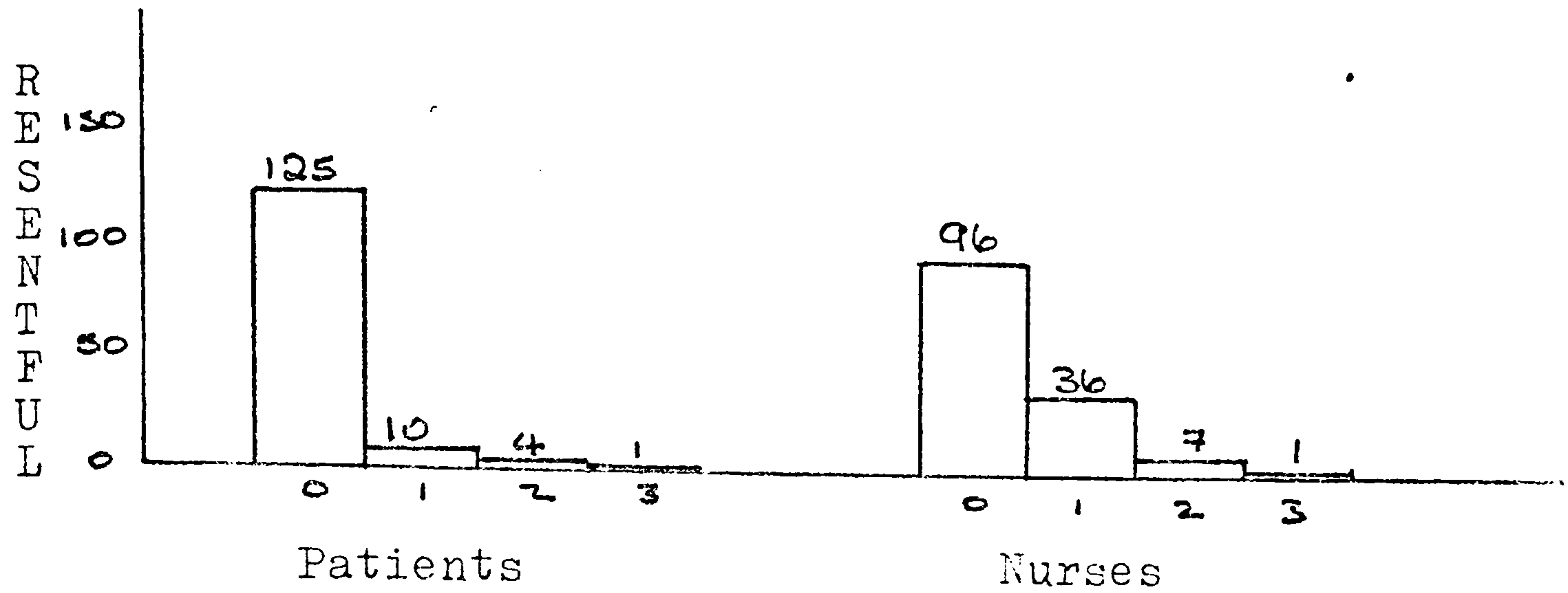


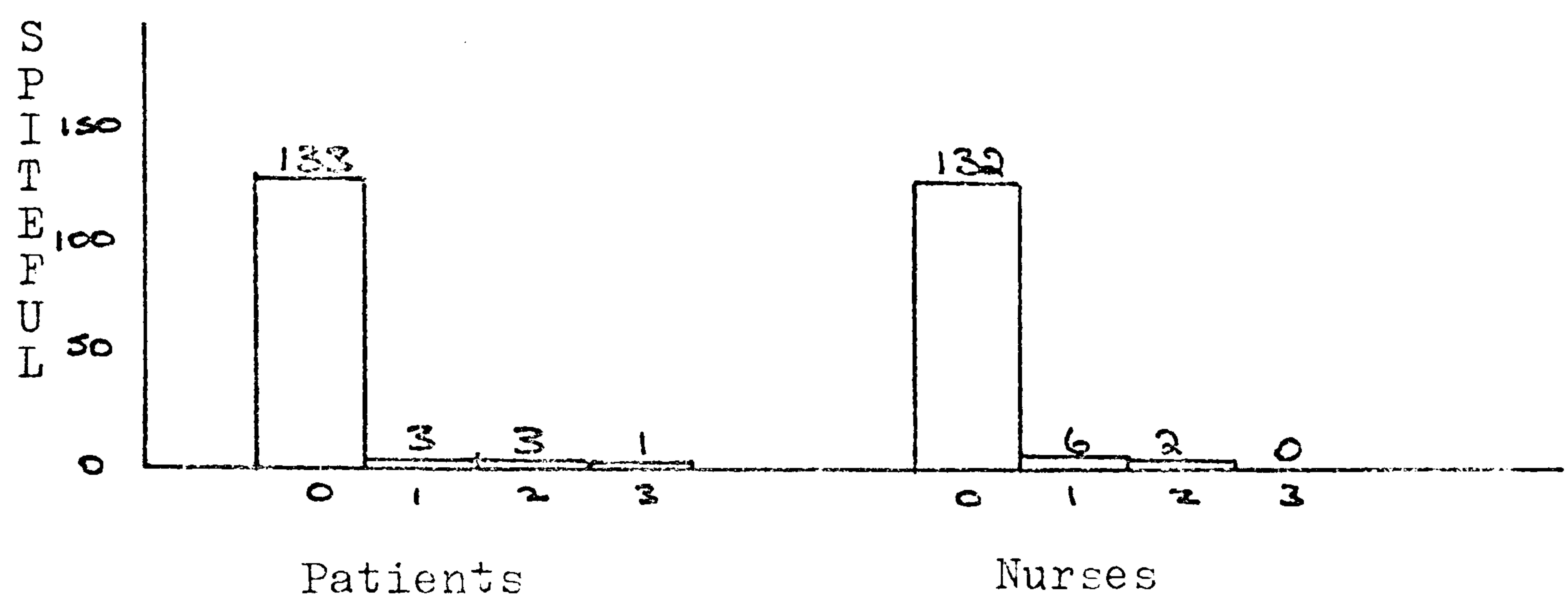


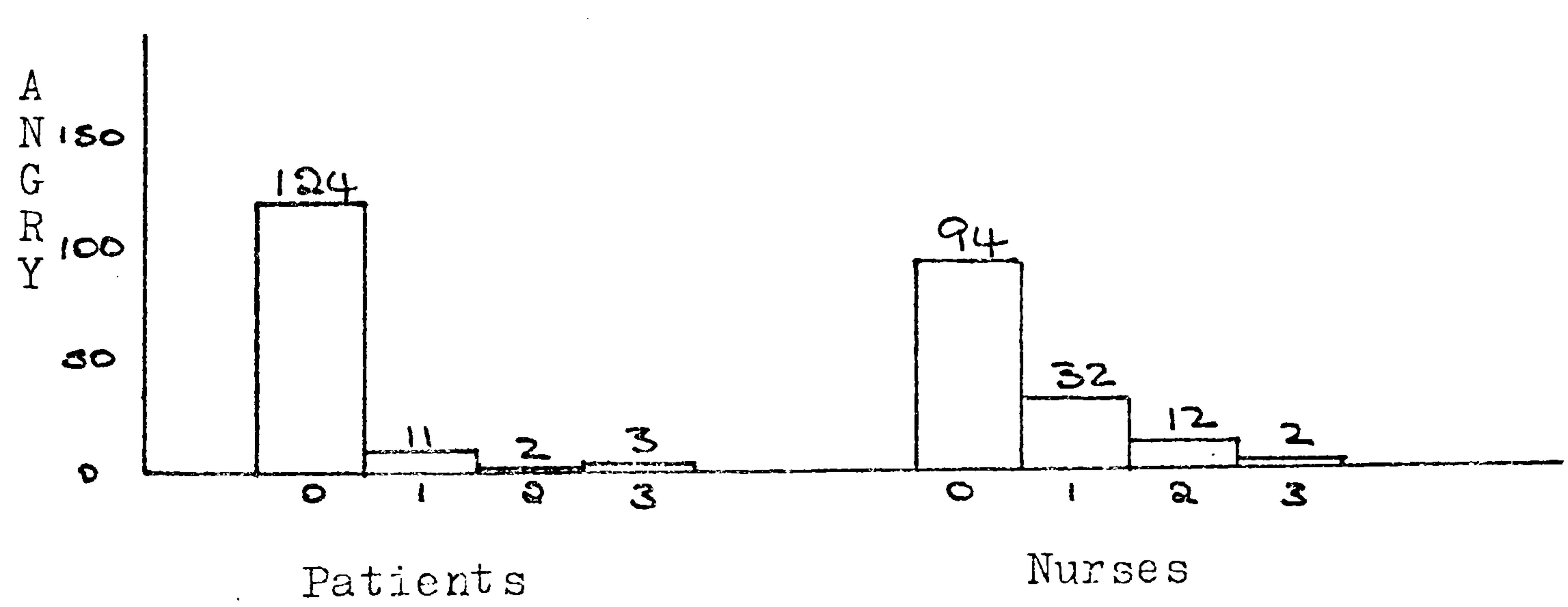


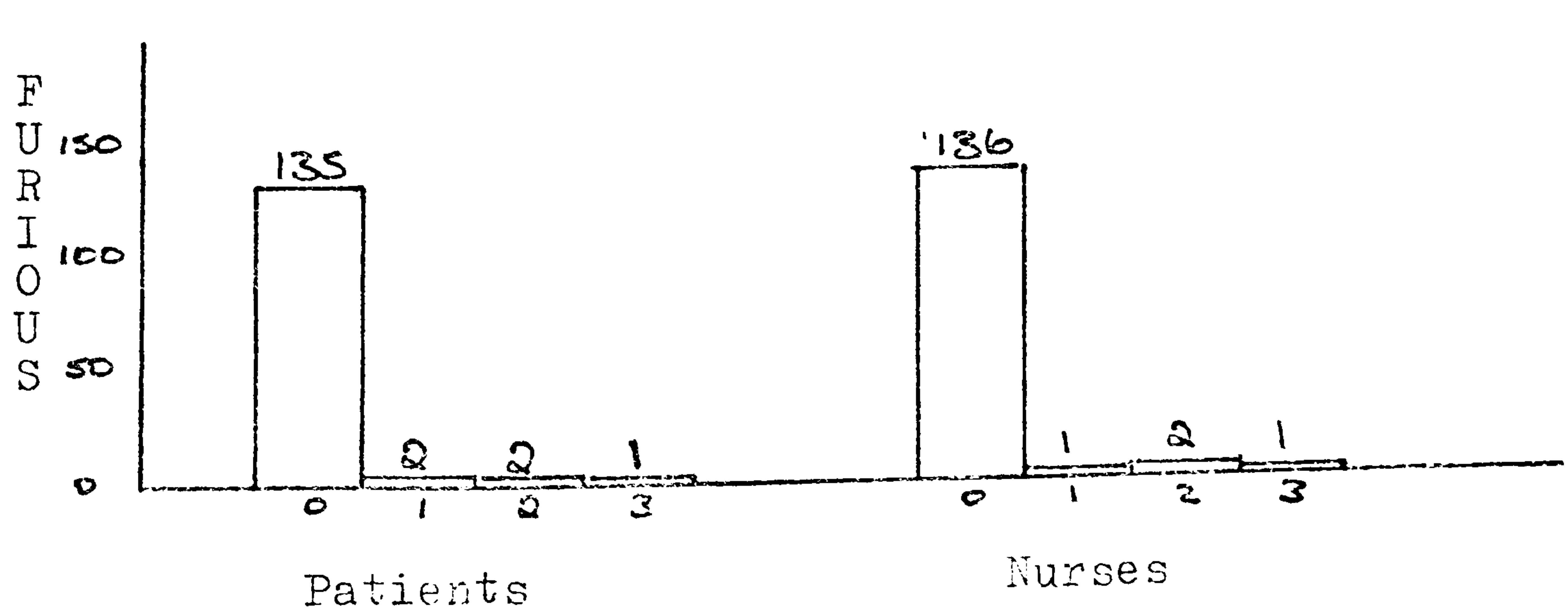


Factor 5 Hostility









APPENDIX 50

Section 3, a-d: Data from the Mood Adjective List
The Factor Analysis

Table 3a: Factors derived from M.A.C.L. data on 202 patients

	Adjectives	Loading	% of variance	Cumulative percentage
1. Vigor	Active	0.74		
	Lively	0.74	23.3	23.3
	Full of pep	0.81		
	Vigorous	0.70		
2. Hostility	Spiteful	0.30		
	Resentful	0.53	11.1	34.4
	Furious	0.43		
	Angry	0.85		
3. Fatigue	Tired	0.60		
	Worn out	0.65	8.8	43.2
	Sluggish	0.65		
	Weary	0.47		
4. Depression	Helpless	0.43		
	Unhappy	0.59		
	Guilty	0.30		
	Hopeless	0.41	6.2	49.4
	Depressed	0.50		
	Miserable	0.65		
	Worthless	0.60		
	Discouraged	0.47		
5. Anxiety	Tense	0.62		
	On edge	0.55	5.3	54.8
	Nervous	0.77		
	Shaky	0.52		

APPENDIX 51

Factor Analysis Patients

Eigen Value	% Var.	Cum. % Var.
2.65013	14	14
2.42644	13	27
1.88589	10	37
1.37857	8	45
1.08660	6	51
1.08309	6	57
1.0306	6	63
.93869	5	68
.91202	5	73
.79364	4	77
.68999	4	81
.67242	4	85
.63799	3	88
.61581	3	91
.48709	2.5	93.5
.29283	1.5	95
.24814	1	96
.20341	1	97
.17453	1	98
.15677	1	99
.04641	.25	99.25
.01	.05	99.3
.02	.01	99.31
.003	.01	99.32

APPENDIX 52

Factor Analysis Nurses

Eigen Value	% Var.	Cum. % Var.
2.84595	15.5	15.4
2.52629	14	29.5
2.40345	13	42.5
1.83349	10	52.5
1.22812	7	59.5
1.20169	6.5	66
.99595	5.5	71.5
.77984	4	75.5
.70886	4	79.5
.6131	3.5	83
.60044	3	86
.59823	3	89
.53969	3	92
.33692	2	94
.2827	1.5	95.5
.22221	1	96.5
.13636	1	97.5
.12755	1	98.5
.12256	1	99.5
.0932	.05	99.55
.07707	.04	99.59
.05546	.03	99.6
.05074	.02	99.6
.01997	.01	99.65

APPENDIX 53

Partial Correlation Coefficients of Nurse/Patient Ratings Controlling for the Effects of Cognitive Complexity

Table A.1

Factor 1 - Vigor

Vigorous	Lively	Full of pep	Active
.2495 $p < .003$	.1499 $p < .078$	.1726 $p < .042$	.1983 $p < .019$

Table A.2

Factor 2 - Depression

Unhappy	Miserable	Discouraged	Depressed
.3210 $p < .000$	.2746 $p < .001$	.1009 $p < .237$	.0666 $p < .436$

Table A.3

Factor 3 - Anxiety

Nervous	Tense	On edge
-.0025 $p < .977$	-.0018 $p < .983$	-.0381 $p < .656$

Table A.4

Factor 4 - Fatigue

Shaky	Weary	Tired	Wornout
.1500 $p < .077$	.1799 $p < .033$	.3419 $p < .000$	.0864 $p < .310$

Table A.5

Factor 5 - Hostility

Resentful	Spiteful	Angry	Furious
.1718 $p < .043$	.0589 $p < .491$	.2774 $p < .001$	.1978 $p < .02$

APPENDIX 54

Correlation Coefficients of Z Score Transformations

Factor 1

	Vigorous	Lively	Full/Pep	Active
Z scores	.2515 p < .003	.1465 p < .084	.1723 p < .042	.1942 p < .021
Raw Data	.2515 p < .003	.1448 p < .088	.1723 p < .042	.2013 p < .017

Factor 2

	Unhappy	Miserbale	Discouraged	Depressed
Z scores	.3210 p < .000	.2857 p < .001	.0945 p < .265	.0698 p < .413
Raw Data	.3210 p < .000	.2812 p < .001	.0992 p < .243	.0683 p < .422

Factor 3

	Nervous	Tense	On edge
Z scores	-.0035 p < .967	-.0102 p < .905	-.0298 p < .73
Raw Data	-.0052 p < .951	-.0096 p < .911	-.0384 p < .652

Factor 4

	Shaky	Weary	Tired	Worn-out
Z scores	.1500 p < .077	.1799 p < .033	.3419 p < .000	.0864 p .310
Raw Data	.1407 p < .097	.1799 p < .033	.3419 p < .000	.0836 p < .33

Factor 5

	Resentful	Spiteful	Angry	Furious
Z scores	.1717 $p < .043$	.0587 $p < .49$	.2449 $p < .004$	.1890 $p < .025$
Raw Data	.1717 $p < .043$	.0587 $p < .49$	.2449 $p < .004$	.1890 $p < .025$

APPENDIX 55

Female/Male Patients

Table A.6

F1 - Vigor

	Female		Male	
	Patient	Nurse	Patient	Nurse ↑
Vigor	–.1194	–.0333	.116	.101 ↑
Lively	–.1843	–.1142	.179	.181 ↑
Pep	–.218	–.068	.21	.117 ↑
Active	–.203	–.217	.198	.246 ↑

Table A.7

F2 - Depression

	Female		Male	
	Patient	Nurse	Patient	Nurse
Unhappy	–.0525	.0244	.05	–.074 ↓
Miserable	.0524	.0581	–.05	–.146 ↓
Discouraged	–.005	–.029	.005	.004 ↓
Depressed	.125	.006	–.122	–.088 ↓

Table A.8

F3 - Anxiety

	Female		Male	
	Patient	Nurse	Patient	Nurse ↓
Nervous	.156	.209	–.15	–.192 ↓
Tense	–.07	.0465	.07	–.152 ↓
On Edge	.1265	.038	–.123	–.083 ↓

Table A.9

F4 - Fatigue

	Female		Male	
	Patient	Nurse	Patient	Nurse ↓
Shaky	-.032	.026	.031	-.078 ↓
Weary	-.012	.074	.01	-.147 ↓
Tired	-.037	.126	.036	-.205 ↓
Worn-out	-.041	.16	.04	-.234 ↓

APPENDIX 56

Stage of Training

Table A.11
F1 - Vigor

	Year 1		Year 3	
	Patient	Nurse	Patient	Nurse
Vigor	-.006	.027	.06	.115
Lively	-.096	.259	1.8	-.196
Pep	.059	.092	-.014	-.036
Active	-.02	.178	7.0	-.01

Table A.12
F2 - Depression

	Year 1		Year 3	
	Patient	Nurse	Patient	Nurse
Unhappy	.149	-.109	-.153	.068
Miserable	.057	-.022	-.099	-.003
Discouraged	.081	.023	-.015	.235
Depressed	.055	-.068	-.102	.09

Table A.13
F3 - Anxiety

	Year 1		Year 3	
	Patient	Nurse	Patient	Nurse
Nervous	-.1315	-.003	-.004	-.044
Tense	.068	-.074	-.098	.059
On Edge	-.046	.006	-.06	.069

Table A.14

F4 - Fatigue

	Year 1		Year 3	
	Patient	Nurse	Patient	Nurse
Shaky	−.046	−.225	.01	.223
Weary	.096	−.018	−.02	.108
Tired	−.042	−.056	.097	.02
Worn-out	−.076	−.048	.0535	.14

APPENDIX 57

Cognitive Complexity

Table A.15

	Low Scores		Mid Scores		High Scores	
	Patient	Nurse	Patient	Nurse	Patient	Nurse
Vigor	.0414	.067	.189	-.073	-.174	-.073
Lively	.23	.181	-.0035	-.0705	.148	-.07
Pep	.092	.172	.112	-.0325	.04	.004
Active	.18	.095	.151	-.041	-.008	-.175

Table A.16

	Low Scores		Mid Scores		High Scores	
	Patient	Nurse	Patient	Nurse	Patient	Nurse
Unhappy	-.094	-.07	.117	-.039	-.008	.1
Miserable	.128	-.202	-.1255	.248	-.026	-.159
Discour.	.056	.017	-.0235	.107	.02	-.164
Depressed	.113	-.112	-.079	.155	-.079	-.004

Table A.17

	Low Scores		Mid Scores		High Scores	
	Patient	Nurse	Patient	Nurse	Patient	Nurse
Nervous	-.0285	-.069	-.123	.034	.062	-.001
Tense	-.156	-.153	-.028	.022	-.133	-.152
On Edge	.18	-.018	-.253	-.15	.035	.097

Table 18

	Low Scores		Mid Scores		High Scores	
	Patient	Nurse	Patient	Nurse	Patient	Nurse
Shaky	.061	-.132	-.269	.008	.103	.076
Weary	.142	-.1785	-.09	.108	.118	-.021
Tired	.166	-.277	-.182	.12	-.09	.053
Worn-out	.150	-.356	-.262	.185	-.035	.019

Table A.19

	Low Scores		Mid Scores		High Scores	
	Patient	Nurse	Patient	Nurse	Patient	Nurse
Resentful	-.075	.210	.035	-.007	-.081	-.051
Angry	-.156	.21	-.211	-.108	.143	-.248

APPENDIX 58

Composition of Sub-Groups of Subjects

	1st year	3rd year	High Score	Medium Score	Low Score
Male Patients	29	31	24	15	16
Female Patients	27	26	14	20	19

	1st year	2nd year	3rd year	Male Pat.	Female Pat.
High Scores	12	11	13	24	14
Medium Scores	18	5	15	15	20
Low Scores	17	2	16	16	19

X^2 Male/Female Distribution as a Function of Cognitive Complexity

	Male	Female	Total
High	24 (19.35)	14 (18.6)	38
Medium	15 (17.8)	20 (17.2)	35
Low	16 (17.8)	19 (17.2)	35
Total	55	53	108

$$X^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

APPENDIX 59

Partial Correlation Coefficients of Nurses Ratings for Patients with Ratings of Assumed Similarity Measure (Controlling Stereotype) and Stereotype Measure (Controlling Assumed Similarity).

N = 140	Assumed Similarity with Judgement of Patient	Stereotype with Judgement of Patient
Vigorous	.299 ($p<.0001$)	.0965 ($p<.258$)
Lively	.419 ($p<.0001$)	.0832 ($p<.330$)
Full/Pep	.2615 ($p<.002$)	.3154 ($p<.0001$)
Active	.426 ($p<.0001$)	.0713 ($p<.404$)
Unhappy	.185 ($p<.03$)	.2785 ($p<.001$)
Miserable	.293 ($p<.0001$)	.222 ($p<.009$)
Discouraged	.293 ($p<.0001$)	.254 ($p<.003$)
Depressed	.277 ($p<.001$)	.246 ($p<.003$)
Nervous	.232 ($p<.006$)	.248 ($p<.003$)
Tense	.150 ($p<.08$)	.398 ($p<.0001$)
On Edge	.139 ($p<.10$)	.3045 ($p<.0001$)
Shaky	.328 ($p<.0001$)	.240 ($p<.004$)
Weary	.423 ($p<.0001$)	.2195 ($p<.009$)
Tired	.391 ($p<.0001$)	.389 ($p<.0001$)
Worn-out	.2725 ($p<.001$)	.392 ($p<.0001$)
Resentful	.08 ($p<.35$)	.243 ($p<.004$)
Angry	.063 ($p<.5$)	.395 ($p<.0001$)

PAGE

NUMBERING

AS ORIGINAL

APPENDIX 60

Partial Correlation Coefficients of Nurse Ratings for Patients with Ratings of Assumed Similarity Measure (Controlling Stereotype) and Stereotype Measure (Controlling Assumed Similarity) as a Function of Stage of Training.

	Assumed Similarity		Stereotype	
	1st Year	3rd Year	1st Year	3rd Year
Vigorous	.348 ($p<.008$)	.2185 ($p<.11$)	.066 ($p<.6$)	.1113 ($p<.4$)
Lively	.353 ($p<.007$)	.445 ($p<.001$)	.0298 ($p<.8$)	.1165 ($p<.4$)
Full/Pep	.439 ($p<.001$)	.32 ($p<.02$)	.024 ($p<.9$)	.630 ($p<.0001$)
Active	.439 ($p<.001$)	.413 ($p<.002$)	.0835 ($p<.5$)	.152 ($p<.26$)
Unhappy	.414 ($p<.001$)	.058 ($p<.67$)	−.105 ($p<.4$)	.428 ($p<.001$)
Miserable	.44 ($p<.001$)	.377 ($p<.004$)	−.054 ($p<.7$)	.2265 ($p<.09$)
Discouraged	.456 ($p<.0001$)	.189 ($p<.16$)	.177 ($p<.19$)	.213 ($p<.115$)
Depressed	.307 ($p<.02$)	.244 ($p<.07$)	.074 ($p<.6$)	.2886 ($p<.03$)
Nervous	.459 ($p<.0001$)	.114 ($p<.4$)	.273 ($p<.04$)	.310 ($p<.02$)
Tense	.202 ($p<.13$)	.100 ($p<.5$)	.372 ($p<.004$)	.185 ($p<.17$)
On Edge	.215 ($p<.11$)	.031 ($p<.8$)	.2905 ($p<.03$)	.234 ($p<.08$)
Shaky	.173 ($p<.2$)	.367 ($p<.005$)	.172 ($p<.2$)	.159 ($p<.2$)
Weary	.546 ($p<.0001$)	.232 ($p<.09$)	.005 ($p<.9$)	.415 ($p<.001$)
Tired	.469 ($p<.0001$)	.293 ($p<.03$)	.065 ($p<.6$)	.139 ($p<.3$)
Worn-out	.348 ($p<.008$)	.396 ($p<.002$)	.247 ($p<.06$)	.244 ($p<.07$)
Resentful	.2098 ($p<.12$)	−.013 ($p<.9$)	.336 ($p<.01$)	.348 ($p<.009$)
Angry	.212 ($p<.11$)	.0638 ($p<.6$)	.2096 ($p<.12$)	.2965 ($p<.03$)

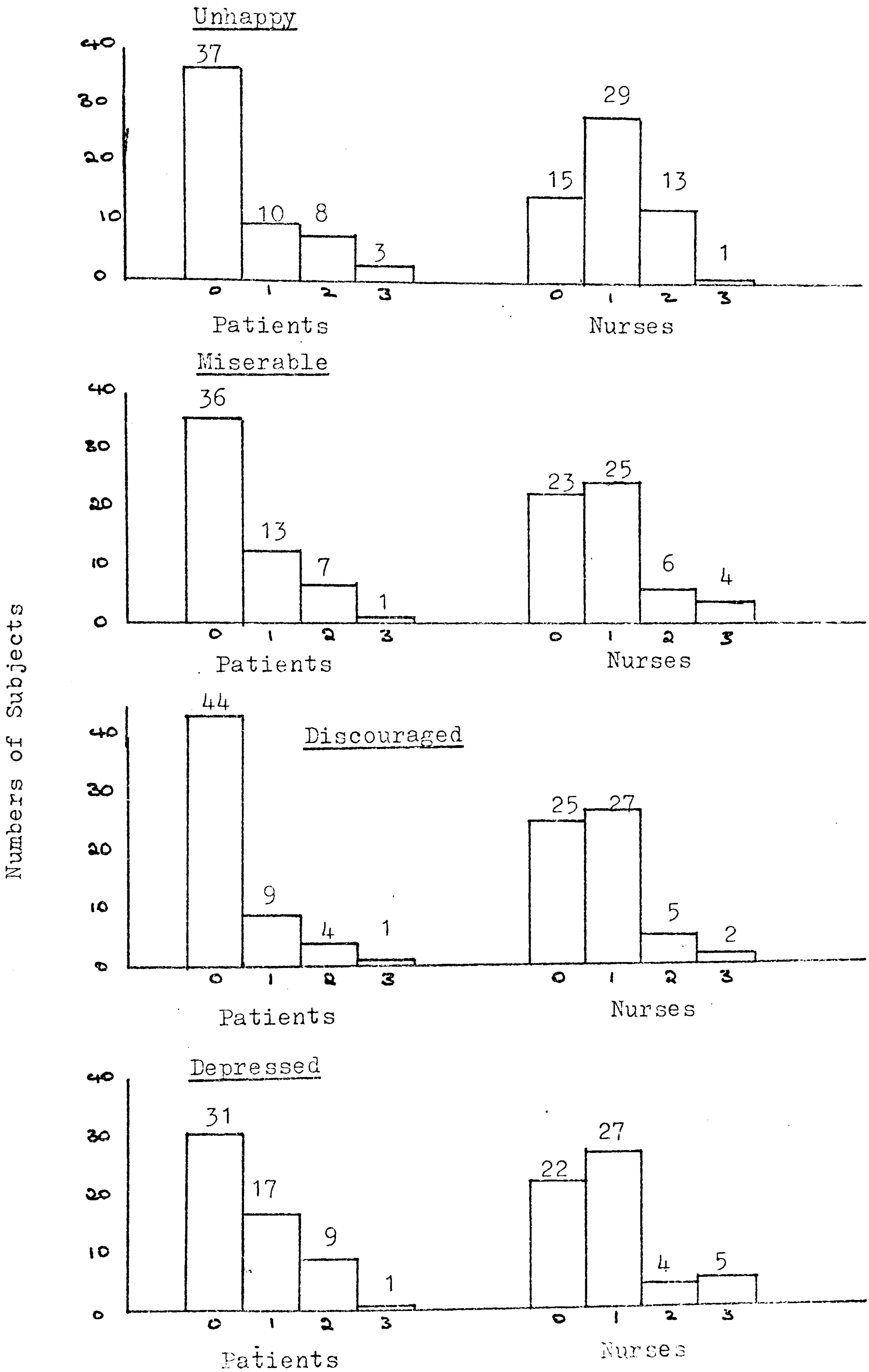
3rd Year N = 57, 1st Year N = 58

APPENDIX 61

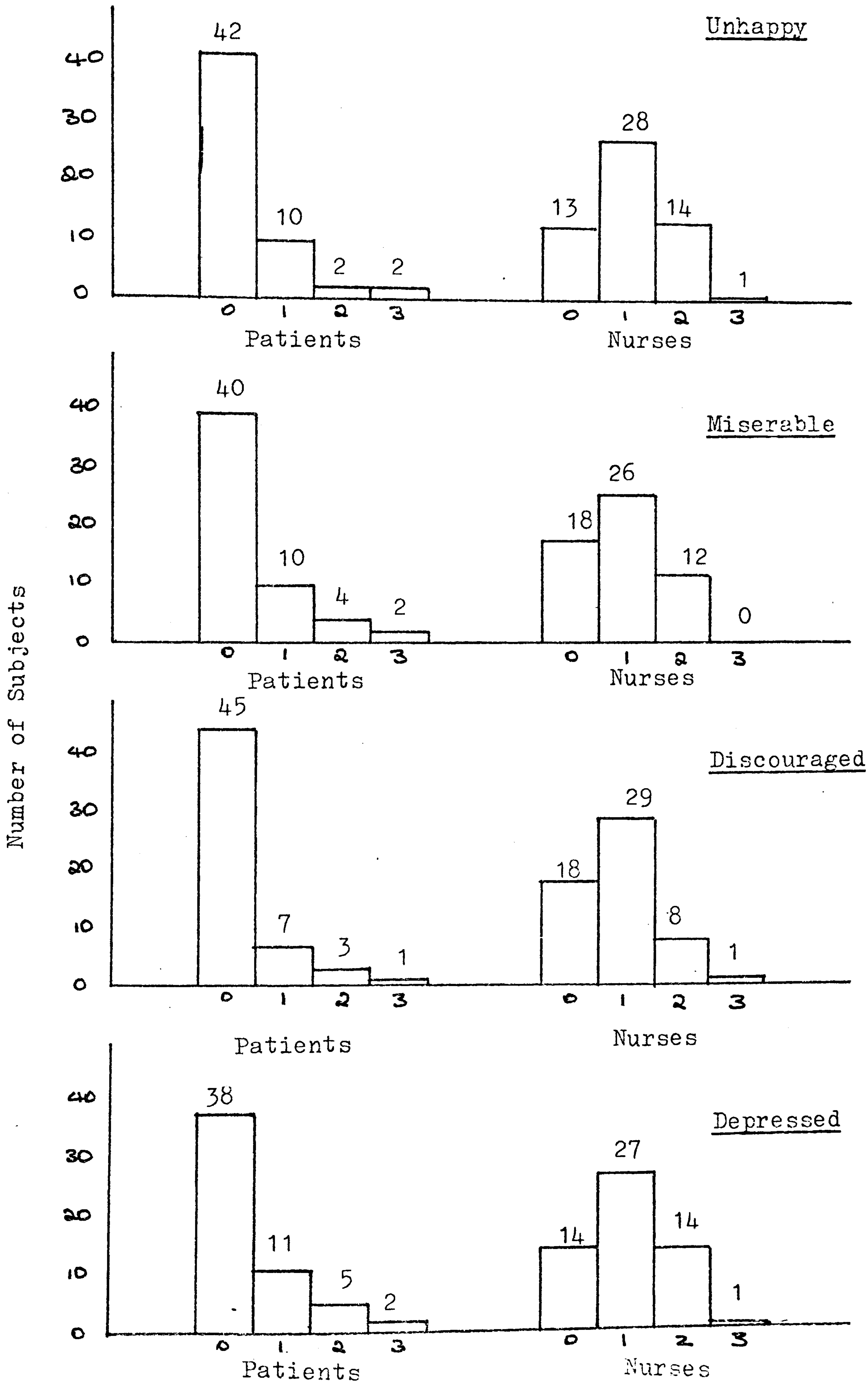
Partial Correlation Coefficients of Nurses Ratings for Patients with Ratings of Assumed Similarity Measure (Controlling Stereotype) and Stereotype Measure (Controlling Assumed Similarity) as a Function of Cognitive Complexity.

	Assumed Similarity		Stereotype	
	High Scores	Low Scores	High Scores	Low Scores
Vigorous	.357 ($p<.35$)	.508 ($p<.002$)	.078 ($p<.6$)	-.106 ($p<.55$)
Lively	.495 ($p<.003$)	.5395 ($p<.001$)	-.082 ($p<.6$)	.0375 ($p<.8$)
Full/Pep	.4775 ($p<.004$)	.009 ($p<.9$)	.0255 ($p<.9$)	.743 ($p<.0001$)
Active	.4455 ($p<.007$)	.1996 ($p<.3$)	.088 ($p<.6$)	.354 ($p<.04$)
Unhappy	.387 ($p<.02$)	.100 ($p<.6$)	.282 ($p<.10$)	.3845 ($p<.025$)
Miserable	.320 ($p<.06$)	.374 ($p<.03$)	.057 ($p<.7$)	.249 ($p<.16$)
Discouraged	.196 ($p<.26$)	.285 ($p<.10$)	.261 ($p<.13$)	.359 ($p<.04$)
Depressed	.517 ($p<.001$)	.303 ($p<.08$)	-.0285 ($p<.9$)	.369 ($p<.03$)
Nervous	.402 ($p<.02$)	.0975 ($p<.6$)	.272 ($p<.11$)	.481 ($p<.004$)
Tense	.124 ($p<.5$)	.084 ($p<.6$)	.304 ($p<.075$)	.367 ($p<.03$)
On Edge	.063 ($p<.7$)	.0855 ($p<.6$)	.375 ($p<.03$)	.380 ($p<.03$)
Shaky	.324 ($p<.05$)	.367 ($p<.033$)	.308 ($p<.07$)	.146 ($p<.4$)
Weary	.197 ($p<.3$)	.499 ($p<.003$)	.564 ($p<.0001$)	.167 ($p<.3$)
Tired	.599 ($p<.0001$)	.422 ($p<.01$)	.054 ($p<.8$)	.288 ($p<.099$)
Worn-out	.456 ($p<.006$)	.215 ($p<.2$)	.198 ($p<.25$)	.595 ($p<.0001$)
Resentful	.213 ($p<.2$)	.108 ($p<.5$)	.168 ($p<.5$)	.223 ($p<.2$)
Angry	.140 ($p<.4$)	.024 ($p<.9$)	.526 ($p<.001$)	.437 ($p<.01$)

Cumulative Distribution of Patients and 1st Year Nurses Rating for the Factor Depression



Frequency Distribution of Patients and 3rd Year Nurses Ratings for the Factor Depression



BIBLIOGRAPHY

Abelson, R.P. (1976) Script Processing Formation and Decision Making. Cognitive and Social Behaviour. Eds. J.S. Carroll & J.W. Payne, New York.

Abelson, R.P. (1981) Psychological Status of the Script Concept. American Psychologist, 36, pp 715-29.

Adams, H.F. (1927) The Good Judge of Personality. Journal Abnormal Social Psychology, 22, pp 172-181.

Adams, G.F. & McIlwraith, P.L. (1963) Geriatric Nursing. London, Oxford University Press.

Adams-Webber, J.R. (1967) Construct and Figure Interactions Within a Personal Construct System: an Extension of Repertory Grid Technique. Unpublished doctoral disertation. Brandeis University.

Adams-Webber, J.R. (1969) Cognitive Complexity and Sociality. British Journal of Clinical & Social Psychology, 8, pp 211-216.

Altschull, A.T. (1970) Go Talk to the Patients. Nursing Mirror, 10th April 130/15, pp 41-46.

Anderson, J. (1978) The Effects of the Patient's Diagnosis on Professionals and Students in a Psychiatric Setting: A Labelling Perspective. J. Nurs. Papers, 10, pp 4-23.

Anderson, J. (1980) Cognitive Units. In Abelson R.P. (1989) 36, pp 715-729. Op. Cit.

Anderson, N.H. (1961) Scales and Statistics: Parametric and Non-Parametric. Psychological Bulletin, 58, pp 305-316.

Anderson, N.H. (1972) Looking for Configurality in Clinical Judgement. Psychological Bulletin, 78, pp 93-102.

Argyle, M., Alkema, F. & Gilmour, R. (1971) The communication of Friendly and Hostile Attitudes by Verbal and Non-Verbal Signals. European Journal of Social Psychology, 3, pp 385-402.

Arkes, H.R. & Harkness, A.R. (1980) Effect of Making a Diagnosis on Subsequent Recognition of Symptoms. Journal Experimental Psychology: Human Learning and Memory.

Atkinson, R.C. (1961) The Observing Response in Discrimination Learning. Journal of Experimental Psychology, 62, pp 253-262.

Atkinson, R.C., Bogartz, W.H. & Turner, R.N. (1959) Supplementary Report: Discrimination Learning with Propabilistic Reinforcement Schedules. Journal of Experimental Psychology, 57, pp 349-350.

Bachrach, H.M. (1976) Empathy: We Know What We Mean, but What Do We Measure? Arch. Gen. Psychiatry, 33, pp 35-38.

Baddeley, A.D. & Hitch, G. (1974) Working Memory. In G.H. Bower (Ed). The Psychology of Learning & Motivation. Vol. 8, pp 47-90, New York. Academic Press.

Barclay, S., Beach, L.R. & Braithwaite, W.P. (1971) Normative Models in the Study of Cognition. Organisational Behavio r & Human Performance, 6, pp 389-413.

Bar-Hillel, M. (1973) On the Subjective Probability of Compound Events. Organisational Behavio r & Human Performance, 9, pp 396-405.

Bar-Hillel, M. (1974) Similarity and Proability. Organisational Behavio r & Human Performance, 11, pp 277-282.

Barrett-Lennard, G.T. (1962) Dimensions of Therapist Response as Causal Factors in Theraputic Change. Psychological Monographs.

Beach, L.R. & Mitchell, T.R. (1978) A Contingency Model for the Selection of Decision Strategies. Academy of Management Review, 3, pp 439-449.

Beck, A.T. (1967) Depression, Clinical, Experimental and Theoretical Aspects. Hoeber Medical Division. Harper and Row.

Beck, J.R. (1983) Microcomputer Aids to Medical Decision Making: Educational and Clinical Applications. Medical Decision Making, 3, pp 3-7.

Becker, J. (1974) Depression: Theory and Research. London, J. Wiley & Sons.

Berkeley, D. & Humphrey, P. (1982) Structuring Decision Problems and the Bias Heuristic. Acta Psychologica, 50, pp 201-252.

Bieri, J., Atkins, A.L., Briar, S., Lobeck, R., Miller, H. & Tripodi, T. (1970) The Nature of Cognitive Structures. In P.B. Warr (Ed). Thought & Personality, Penguin Books.

Bond, S. (1978) Processes of Communication About Cancer in a Radiotherapy Department. Unpublished PhD Thesis: University of Edinburgh.

Boneau, C.A. (1960) The Effects of Violations of Assumptions Underlying the 't' Test. Psychological Bulletin, 57, pp 49-64.

Bower, G., Black, J. & Turner, T. (1979) Scripts in Text Comprehension and Memory. Cognitive Psychology, 11, pp 177-220.

Bowling, A. (1981) Delegation in General Practice. A Study of Doctors and Nurses. London, Tavistock Publications Limited.

Brady, M.M. (1976) Nurses Attitudes Towards a Patient who has a Psychiatric History. Journal Advanced Nursing, 1, pp 11-23.

Bransford, J.D., Barclay, J.R. & Franks, J.J. (1972) Sentence Memory: A Constructive vs Interpretive Approach. Cognitive Psychology, 3, pp 193-209.

Brown, G.W. (1972) Life Events and Psychiatric Illness: Some Thoughts on Methodology and Causality. Journal of Psychosomatic Research, 16, pp 311-320.

Bruner, J.S., Goodnow, J.J. & Austin, G.A. (1956) A Study of Thinking. New York, Wiley.

Buck, R., Baron, R., Goodman, N. & Shapiro, B. (1980) Unitization of Spontaneous Non-verbal Behaviour in the Study of Emotion Communication. J. Personality Social Psychology, 39, pp 522-29.

Burke, C.J. & Estes, W.K. (1957) A Component Model for Stimulus Variables in Discrimination Learning. Psychometrika, 22, pp 133-145.

Burns, K.L. & Beier, E.G. (1973) Significance of Vocal and Visual Channels in Decoding of Emotional Meaning. Journal of Communication, 23, pp 118-130.

Cantor, N. & Mischell, W. (1977) Traits as Prototypes: Effects on Recognition Memory. J. Personality and Social Psychology, 35, pp 38-48.

Carnevali, D.L. (1966) Preoperative Anxiety. American Journal of Nursing, 7, pp 1536-1538.

Cartwright, A. (1964) Human Relations and Hospital Care. Institute of Community Studies. London, Routledge Kegan Paul.

Castallan, N.J. (1973) Multiple-Cue Probability Learning with Irrelevant Cues. Organizational Behavio r and Human Performance, 9, pp 16-29.

Castallan, N.J. (1977) Decision Making with Multiple Probabilistic Cues. In N.J. Castallan., D.B. Pisoni & G.R. Potts (Eds). Cognitive Theory (Vol 2). Hillside N.J. Erlbaum.

Castallan, N.J. & Edgell, S.E. (1973) An Hypothesis Generation Model in Non-Metric Multiple-Cue Probability Learning. Journal of Mathmatical Psychology, 10, pp 204-222.

Castelnuovo-Tedescu, P. (1961) Depression in Patients with Physical Disease. Cranbury. Wallace Labs. New Jersey.

Christiansen-Szalanski, J.J.J. (1978) Proble,-Solving Strategies: A Selection Mechanism, Some Implications and Some Data. Organization Behavio r & Human Performance, 22, pp 307-323.

Christiansen-Szalanski, J.J.J. (1980) A Further Examination of the Selection of Problem-Solving Strategies: The Effects of Deadlines and Analytic Aptitudes. Organizational Behavio r & Human Performance, 25, pp 107-122.

Christiansesn-Szalanski, J.J.J. & Beach, L.R. (1982) Experience and the Base-Rate Fallacy. Org. Behav. & Hu. Perfor. 29, pp 270-278.

Christiansen-Szalanski J.J.J. & Bushyhead, J.B. (1981) Physicians' use of Probabilistic Information in a Real Clinical Setting. Journal of Experimental Psychology: Human Perception & Performance, 7, pp 928-935.

Clarke, J.R. (1982) The Role of Decision Skills and Medical Knowledge in the Clinical Judgement of Surgical Residents. Surgery, 92, pp 153-158.

Cline, U.B. (1964) Interpersonal Perception. _Progress in Experimental Personality Research_, Vol. 1, pp 183-192.

Cohen, L.J. (1979) On the Psychology of Prediction: Whose is the Fallacy? _Cognition_, 7, pp 385-407.

Cook, M. (1971) Interpersonal Perception. _Penguin Books_.

Cronbach, L.J. (1955) Processes Affecting Scores on Understanding of Others and Assumed Similarity. _Psychological Bulletin_, 52, pp 177-193.

Darwin, C. (1972) The Expression of the Emotions in Man and Animals. London: John Murray.

Dawes, R.M. (1976) Shallow Psychology. _Cognition and Social Behaviour_. Eds. J.S. Carroll & J.W. Payne, New York, Lawrence Erlbaum.

Davitz, J.R. & Davitz, L.L. (1981) Inferences of Patients' Pain and Psychological Distress. New York, Springer.

den Uyl, M. & van Oostendorp, H. (1980) The use of Scripts in Text Comprehension. _Poetics_, 9, pp 275-294.

De Paulo, B., Rosenthal, R., Eisenstat, R., Finkelstein, S. & Rogers, P. (1978) Decoding Discrepant Non-Verbal Cues. _Journal of Personality & Social Psychology_, 36, pp 313-323

De Wolfe, A.S., Barrell, R.P. & Cummings, J.W. (1966) Patient Variables in Emotional Response to Hospitalisation for Physical Illness. _Journal of Consulting Psychology_, 30, pp 68-72.

Edwards, A.L. (1957) The Social Desirability Variable in Personality Assessment and Research. New York, Dryden Press.

Einhorn, H.J. (1972) Expert Measurement and Mechanical Combination. Organizational Behavior & Human Performance, 7, pp 86-106.

Einhorn, H.J. (1974) Expert Judgement: Some Necessary Conditions and an Example. Journal Applied Psychology, 59, pp 562-571.

Einhorn, H.J. & Hogarth, R.M. (1978) Confidence in Judgement: Persistences of the Illusion and Validity. Psychological Review, 85, pp 395-416.

Einhorn, H.J. & Hogarth, R.M. (1981) Behavioural Decision Theory, Processes of Judgement and Choice. Annual Review of Psychology, 32, pp 53-88.

Eisenberg, J.M. (1979) Sociologic Influences on Decision-Making by Clinicians. Annals of Internal Medicine, 90, pp 957-964.

Ekman, P. (1973) Cross-Cultural Studies of Facial Expression. In P. Ekman (Ed). Darwin and Facial Expression: A Century of Research in Review, pp 169-222. New York, Academic Press.

Ekman, P. (1979) About Brows: Emotional and Conversational Signals. In. M. von Cranach, K. Foppa, W. Lepenies & D. Ploog (Eds) Human Ethology. Cambridge University Press.

Ekman, P., Friesen, W.V., O'Sullivan, M. & Scherer, K. (1980) Relative Importance of Face, Body and Speech in Judgements of Personality and Affect. Journal of Personality & Social Psychology, 38, pp 270-277.

Ekman, P., Hagar, J.C. & Friesen, W.V. (1981) The Symmetry and Emotional and Deliberate Facial Action. Psychophysiology, 18, pp 101-106.

Ekman, P. & Oster, H. (1982) Review of Research, 1970-1980. Emotion in the Human Face, 2nd Ed. pp 147-173. Cambridge University Press.

Ekman, P., Sorenson, E.R. & Friesen, W.V. (1969) Pan-Cultural Elements in Facial Displays of Emotions. Science, 164, pp 86-88.

Elstein, A.S. & Bordage, G. (1979) Psychology of Clinical Reasoning. Health Psychology, Eds. G. Stone, F. Cohen & N. Adla. San Francisco: Jossey-Bass. ✗

Elstein, A.S., Holmes, M.M., Ravitch, M.M., Rovner, P.R., Holzman, G.B. & Rothert, M.L. (1983) Medical Decisions in Perspective: Applied Research in Cognitive Psychology. Perspectives in Biology & Medicine, 26, pp 486-501. X

Elstein, A.S., Shulman, L.S. & Sprafka, S.A. (1978) Medical Problem Solving: An Analysis of Clinical Reasoning. Cambridge, Mass, Harvard University Press. Cited in Elstein A.S. et al. (1983).

Estes, W.K. (1976) The Cognitive Side of Probability Learning. Psychological Review, 83, pp 37-64.

Eysenck, H.J. ' Eysenck, S.B.G. (1964) Manual of the Eysenck Personality Inventory. University of London Press.

Fischoff, B., Slovic, P. & Lichstein, S. (1978) Fault Trees; Sensitivity of Estimated Failure Probabilities to Problem Representation. Journal of Experimental Psychology: Human Perception & Performance, 4, pp 330-344.

Forsyth, G.L. (1979) Exploration of Empathy in Nurse-Client Interaction. Advanced Nursing Science, 1, pp 53-61.

Fream, W.C. (1978) Notes of Surgical Nursing, 2nd Ed. London: Churchill Livingstone.

Friesen, W.V. (1972) Cultural Differences in Facial Expression in a Social Situation: An Experiemntal Test of the Concept of Display Rules. Unpublished doctoral dissertion. University of California, San Francisco.

Frijda, N.H. (1969) Recognition of Emotion. Advances in Experimental Social Psychology (Vol. 4), pp 167-223. New York, Academic Press.

Funaro, J.F. (1975) An Empirical Analysis of Five Descriptive Models for Cascade Inference. Organizational Behavio r & Human Performance, 14, pp 186-206.

Gettys, C.F., Kelly, C.W. & Peterson, C.R. (1973) The Best Guess Hypothesis in Multistage Inference. Organ. Behav. & Human Per. 10, pp 364-373.

Gillmore, M.R. & Hill, C.T. (1981) Reactions to Patients who Complain of Pain: Effects of Ambiguous Diagnosis. J. of Applied Social Psychology, 11, pp 14-22.

Goddard, H. (1953) The Work of Nurses in Hospital Wards. Oxford, Nuffield Provincial Hospitals Trust.

Goldberg, L.R. (1968) Simple Models or Simple Processes? Some Research on Clinical Judgement. American Psychologist, 23, pp 483-496.

Gordon, M. (1980) Predictive Strategies in Diagnostic Tasks. Nursing Research, 29, pp 39-45.

Gottheil, E., Exline, R.V. & Winkelmayer, R. (1979) Judging Emotions of Normal and Schizophrenic Subjects. American Journal Psychiatry, 136, pp 1049-1054.

Graesser, A.C., Woll, S.B., Kowalski, D.J. & Smith, D.H. (1980) Memory for Typical and Atypical Actions in Scripted Activities. J. of Exp. Psychol: Human Learning & Memory, 6, pp 503-515.

Grosz, H.J. & Grossman, G.K. (1968) Clinicians Response Style: A Source of Variation and Bias in Clinical Judgements. Journal Abnormal Psychology, 73, pp 207-214.

Hamilton, D.L. (1976) Cognitive Biases in the Perception of Social Groups. J.S. Carroll & J.W. Payne (Eds) Cognition and Social Behaviour. Hillside, New Jersey, Erlbaum.

Hamilton, D.L. (1979) A Cognitive-Attributional Analysis of Stereotyping. L. Berkowitz (Ed). Advances in Experimental Psychology, Vol. 12. New York, Academic Press.

Hamilton, D.L. & Rose, T.L. (1980) Illusory Correlation and the Maintenance of Stereotypic Beliefs. Journal Personality & Social Psychology, 39, pp 832-845.

Hammerton, M. (1973) A Case of Radical Probability Estimation. Journal of Experimental Psychology, 101, pp 252-254.

Hammond, L.K. (1970) Cognitive Structure and Clinical Inference. Journal of General Psychology, 83, 107-118.

Hammond, K.R. (1978) Toward Increasing Competence of Thought in Public Policy Formation. K.R. Hammond (Ed), Judgement & Decision in Public Policy Formation. Denver, Westview Press.

Hammond, K.R., Kelly, K.J., Castellan, N.J., Schneider, R.J. & Vancini, M. (1966) Clinical Inference in Nursing: Use of Information-Seeking Strategies by Nurses. Nursing Research, 15, pp 330-336.

Hammond, K.R., Kelly, K.J., Schneider, R.J. & Vancini, M. (1966) Clinical Inference in Nursing: Information Units Used. Nursing Research, 15, pp 236-243.

Hammond, K.R., Kelly, K.J., Schneider, R.J. & Vancini, M. (1967) Clinical Inference in Nursing: Revising Judgements. Nursing Research, 16, pp 38-45.

Harding, C.M., Eiser, J.R. & Kristiansen, C.M. (1982) The Representation of Mortality Statistics and the Perceived Importance of Causes of Death. Journal Applied Social Psychology, 12, pp 169-181.

Hayward, J. (1975) Information - A Prescription Against Pain. The Study of Nursing Care. R.C.N. London.

Higgins, P. (1980) Comments on Watts (1980); Clinical Judgement and Clinical Training: An Editorial Contribution: Decision Making in General Practice. British Journal of Medical Psychology, 53, pp 193-195.

Hochberg, J.F.P. & Brooks, V. (1977) Motion Organization in Stop Action Sequences. Scandinavian Journal Psychology, 18, pp 187-191.

Hogan, R. (1969) Development of an Empathy Scale. Journal Consulting Clinical Psychology, 33, pp 307-316.

Hogarth, R.M. & Makridakis, S. (1981) Forecasting & Planning: An Evaluation. Management Science, 27, pp. 115-138

Holt, R.R. (1958) Yet Another Look at Clinical and Statistical Prediction: Or is Clinical Psychology Worthwhile? American Psychologist, 25, pp 337-349.

Hsu, T.C. & Feldt, L.S. (1969) The Effect of Limitation on the Number of Criterion Score Values on Significance Level of the F. Test. American Educational Research Journal, pp 515-527.

Hugh Jones, P., Tanser, A.R. & Whitby, C. (1964) Patients View of Admission to a London Teaching Hospital. British Medical Journal, 12th Sept. pp 660-664.

Humphreys, P.C. & McFadden, W. (1980) Experiences with MAUD; Aiding Decision Structuring Versus Bootstrapping the Decision Maker. _Acta Psychologica_, 45, pp 51-69.

Izard, C.E. (1971) The Face of Emotion. New York, Appleton-Century-Crofts.

Janis, I.L. & Mann, L. (1977) Decision Making: A Psychological Analysis of Conflict, Choice and Commitment. New York, _The Free Press_.

Jenkins, H.M. & Ward, W.C. (1965) Judgement of Contingency Between Responses and Outcomes. _Psychological Mono-Graphs_: General & Applied 79 (1 whole No, 594).

Johnson, P.E., Hassebrock, F., Durán, A.S. & Moller, J.H. (1982) Multimethod Study of Clinical Judgement. _Organizational Behavio r & Human Performance_, 30, pp 201-230.

Johnston, M. (1976) Communication of Patients' Feelings in Hospital. Bennett, A.E. (Ed). _Communication Between Doctors and Patients_. London: Nuffield Provincial Hospitals Trust.

Johnston, M. (1980) Anxiety in Surgical Patients. _Psychological Medicine_, 10, pp 145-152.

Johnston, M. (1982) Recognition of Patients Worries by Nurses and by Other Patients. Brit. J. Clin. Psychol. 21, pp 255-261.

Kahneman, D. (1973) Attention and Effort. Englewood Cliffs, N.J. Prentice-Hall.

Kahneman, D. & Tversky, A. (1973) On the Psychology of Pre-diction. _Psychological Review_, 80, pp 237-281.

Kahneman, D. & Tversky, A. (1982a) On the Study of Statistical Intuitions. _Cognition_, 11, pp 123-141.

Kahneman, D. & Tversky, A. (1982b) Variants of Uncertainty, Cognition, 11, pp 143-157.

Kassirer, J.P. & Gorry, G.A. (1978) Clinical Problem Solving: A Behavioural Analysis. Annals Internal Medicine, 89, pp 245-255.

Kelly, G.A. (1955) The Psychology of Personal Constructs, Vol. 1. New York, Norton.

Kendall, P.C. & Watson, D. (1981) Psychological Preparation of Stressful Medical Procedures. In C. Prokop & L. Bradley (Eds), Medical Psychology: Contributions to Medical Psychology. London, Academic Press.

Kendell, R.D. (1976) The Classification of Depressions: A review of Contemporary Confusion. British Journal Psychiatry, 129, pp 15-28.

Kleiter, G.D., Gashowetz, H. & Huber, D. (1976) Bibliography: Decision Making. Salzburg, Austria: Psychological Institute University, Salzburg.

Landis, C. (1924) Studies of Emotional Reactions: II. General Behaviour and Facial Expression. Journal Comparative Psychology, 4, pp 447-509.

Landis, C. (1929) The Interpretation of Facial Expression in Emotion. Journal General Psychology, 2, pp 59-72.

Langer, E.J. (1975) The Illusion of Control. Journal Personality Social Psychology, 32, pp 311-328.

Lau, J., Kassirer, J.P. & Pauker, S.G. (1983) Decision Maker 3.0: Improved Decision Analysis by Personal Computer. Medical Decision Making, 3, pp 39-43.

Leventhal, H. (1957) Cognitive Processes and Interpersonal Predictions. Journal Abnormal Social Psychology, 55, pp 176-180.

Lipowski, Z.J. (1967) Review of Consultation Psychiatry and Psychosomatic Medicine. II, Clinical Aspects. Psychosomatic Medicine, 29, pp 201-224.

Lishman, W.A. (1972) Selective Factors in Memory, Part 2. Affective Disorders. Psychological Medicine, 2, pp 248-253.

Lord, F.M. (1953) On the Statistical Treatment of Football Numbers. American Psychologist, 8, pp 750-751.

Lyon, D. & Slovic, P. (1976) Dominance of Accuracy Information and Neglect of Base Rates in Probability Estimation. Acta Psychologica, 40, pp 286-298.

Macleod Clark, J. (1983) Nurse-Patient Communication. An Analysis of Conversations from Surgical Wards. Nursing Research: Ten Studies in Patient Care. Ed. J. Wilson-Barnett, Chichester: J. Wiley.

Maguire, G.B., Julier, D.L., Hanson, K.E. & Bancroft, J.H.J. (1974) Psychiatric Morbidity and Referral in Two General Medical Wards. British Medical Journal, 16th Feb. pp 268-270.

Manchester Regional Hospital Board and University of Manchester (1955) The Work of the Mental Nurse, Manchester University Press.

Markus, H. (1977) Self-Schemata and Processing Information About the Self. Journal Personality Social Psychology, 35, pp 63-78.

McGilloway, F.A. (1980) The Nursing Process: A Problem-Solving Approach to Patient Care. International Journal Nursing Studies, 17, pp 79-90.

McNair, D.M. & Lorr, M. (1964) An analysis of Mood in Neurotics. Journal Abnormal Social Psychology, 69, pp 620-627.

McWhinney, I.R. (1972) Problem-solving & Decision-making in Primary Medical Practice. Proceedings Royal Society Medicine, 65, pp 934-938

Medin, D.L., Altom, M.W., Edelson, S.M. & Freko, D. (1982) Correlated Symptoms and Simulated Medical Classification. Journal Experimental Psychology: Learning, Memory and Cognition, 8, pp 37-50.

Meehl, P.E. (1954) Clinical Versus Statistical Predictions. Minneapolis: University of Minneapolis Press.

Mehrabian, A. & Ferris, S. (1967) Inference of Attitudes from Non-Verbal Communication in Two Channels. Journal Consulting Psychology, 31, pp 248-252.

Mehrabian, A. & Wiener, M. (1967) Decoding of Inconsistant Communication. Journal Personality Social Psychology, 6, pp 108-114.

Miller, I.W. & Norman, W.H. (1979) Learned Helplessness in Humans: A Review and Attribution-Theory Model. Psychological Bulletin, 86, pp 93-118.

Miller, S.M. (1979) Controllability and Human Stress: Methods, Evidence and Theory. Behavioral Research, 17, pp 287-304.

Moffic, H.S. & Paykel, E.S. (1975) Depression in Medical Inpatients. British Journal Psychiatry, 126, pp 346-353.

Morlock, H. (1967) The Effect of Outcome Desirability on Information Required for Decisions. Behavioral Science, 12, pp 296-300.

Morrow, G.R., Craytor, J.K., Brown, J. & Fass, M. (1976) Nurses Perceptions of Themselves, Cancer Nurses, Typical, Ideal and Cancer Patients. Percept. & Motor Skills, 43, pp 1083-1091.

Newell, A. & Simon, H.L. (1972) Human Problem Solving. Englewood Cliffs, Prentice Hall.

Nisbett, R.E. & Borgida, E. (1975) Attribution and the Psychology of Prediction. Journal Personality Social Psychology, 32, pp 932-943.

Olson, C.L. (1976) Some Apparent Violations of the Representativeness Heuristic in Human Judgement. Journal Experimental Psychology, 2, pp 599-608.

Oskamp, S. (1965) Overconfidence in Case-Study Judgements. Journal Consulting Psychology, 29, pp 261-265.

Payne, J.W. (1982) Contingent Decision Behaviour. Psychological Bulletin, 92, pp 382-402.

Pepper, R. (1977) Professionalism, Training and Work. A Study of Nursing in a General Hospital. Unpublished PhD Thesis, University of Kent.

Peterson, C.R. (1973) Special Issue: Cascade Inference. Organizational Behavior & Human Performance, 10, pp 315-432.

Peterson, C.R. & Beach, L.R. (1967) Man as an Intuitive Statistician. Psychological Bulletin, 68, pp 29-46.

Pitman, E.J.G. (1937) Significance Tests which may be Applied to Samples from any Populations. Supplement to Journal Royal Statistical Society, 4, pp 119-130.

Pollard, P. (1982) Human Reasoning: Some Possible Effects of Availability. Cognition, 12, pp 65-96.

Richards, J.M. (1963) Reconceptualization of the Clinical and Statistical Prediction Controversy in Terms of Components of Accuracy of Interpersonal Perception Scores. Psychological Reports, 12, pp 443-448.

Ridgeway, V. & Mathews, A. (1982) Psychological Preparation for Surgery: A Comparison of Methods. British Journal Clinical Psychology, 21, pp 271-280.

Rosenkrantz, P., Vogel, S., Bee, H., Broverman, I. & Broverman, D. (1968) Sex-Role Stereotypes and Self-Concepts in College Students. Journal Consulting Clinical Psychology, 32, pp 287-295.

Russo, J.E. (1977) The Value of Unit Price Information. Journal Market Research, 14, pp 193-201.

Schiffenbauer, A. (1974) Effects of Observers Emotional State on Judgements of the Emotional State of Others. Journal Personality Social Psychology, 30, pp 31-35.

Schneider, D.J. (1973) Implicit Personality Theory: A Review. Psychological Bulletin, 79, pp 294-319.

Schneider, D.F., Hastord, A.H. & Ellsworth, P.C. (1979) Accuracy in Person Perception. Person Perception, 2nd Ed. Reading Massachussetts: Addison Wesley, pp 204-223.

Schwab, J.J., Bialow, M., Brown, J.M. & Holzer, C.E. (1967) Diagnosing Depression in Medical In-Patients. Annals Internal Medicine, 67, pp 695-707.

Selfridge, O.G. (1959) Pandemonium: A Paradigm for Learning. Mechanisation of Thought Processes. London: HMSO 99, 103.

Seligman, M.E.P. (1978) Comment and Integration. Journal Abnormal Psychology, 17, pp 165-179.

Shanteau, J. (1972) Descriptive Versus Normative Models of Sequential Inference Judgement. Journal Experimental Psychology, 93, pp 63-68.

Shanteau, J. (1975) An Information-Integration Analysis of Risky Decision Making. M.F. Kaplan & S. Schwartz (Eds), Human Judgement & Decision Processes. New York, Academic pp 110-134.

Sherman, M. (1927) The Differentiation of Emotional Responses in Infants, II, the Ability of Observers to Judge the Emotional Characteristics of the Crying of Infants and the Voice of an Adult. Journal Comparative Psychology, 7, pp 335-351.

Shortliffe, E.H. (1976) Computer-Based Medical Consultations: MYCIN. Artificial Intelligence Series. American Elsevier Publishing Co, Inc.

Shortliffe, E.H., Buchanan, B.G. & Feigenbaum, E.A. (1979) Knowledge Engineering for Medical Decision Making: A Review of Computer-Based Clinical Decision Aids. Proceedings IEEE, 67, pp 1207-1224.

Siegel, S. (1956) Non-Parametric Statistics for the Behavioural Sciences. Tokyo, McGraw Hill Kogakiesha Ltd.

Silverstein, M.D. (1983) A Clinical Decision Analysis Program for the Apple Computer. Medical Decision Making, 3, pp 29-37.

Simon, H.A. (1955) A Behavioural Model of Rational Choice. Quart.' Journal Economics, 69, pp 99-118.

Simon, H.A. (1979) Rational Decision Making in Business Organizations. American Economic Review, 69, pp 493-513.

Slovic, P., Fischoff, B. & Lichtenstein, S. (1976) Cognitive Processes and Societal Risk Taking. In Cognition & Social Behaviour. Eds J.S. Carrol & J.W. Payne, Hillside, N.J. Erlbaum.

Slovic, P., Fischoff, B. & Lichtenstein, S. (1977) Behavioural Decision Theory. Annual Review Psychology, 28, pp 1-39.

Slovic, P. & Lichtenstein, S. (1971) Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgement. Organizational Behavior Human Performance , 6, pp 649-740.

Smedslund, J. (1963) The Concept of Correlation in Adults. Scandinavian Journal Psychology, 4, pp 165-171.

Smedslund, J. (1966) Note on Learning, Contingency and Clinical Experience. Scan. J. Psychol. 7, pp 265-266.

Smith, A.E. (1957) Rapport and Social Perception in Mother Adolescent Relationships. Unpublished Doctoral Dissertation. University of Illinois, Chicago.

Snapper, K.J. & Fryback, D.G. (1971) Inferences Based on Unreliable Reports. Journal Experimental Psychology, 87, pp 401-404.

Sniezek, J.A. (1980) Judgements of Probabilistic Events: Remembering the Past and Predicting the Future. Journal Experimental Psychology: Human Perception Performance.

Snyder, M., Tanke, E.D. & Berscheid, E. (1977) Social Perception and Interpersonal Behaviour. On the Self-Fulfilling Nature of Scoial Stereotypes. Journal Personality Social Psychology, 35, pp 656-666.

Spielberger, C.D. (1972) Anxiety, Current Trends in Theory and Research. Vol. I. London, Academic Press.

Steiger, J.H. & Gettys, C.F. (1972) Best-Guess Errors in Multistage Inference. J. Exper. Psychol. 92, pp 1-7.

Stockwell, F. (1975) The Unpopular Patient. The Study of Nursing Care. R.C.N.

Sulin, R.A. & Dooling, D.J. (1974) Intrusion of Thematic Idea in Retention of Prose. J. Exper. Psychol. 103, pp 255-262.

Summers, S.A. (1968) Alternative Bases for Choice in Probabilistic Discrimination. J. Exper. Psychol. 76, pp 538-543.

Sutherland, N.F. (1973) Object Recognition. In E.C. Carterette & M.P. Friedan (Eds) Handbook of Perception Vol III. New York, Academic Press.

Tajfel, H. (1978) Intergroup Behaviour: II Group Perspectives. H. Tajfel & C. Fraser (Eds) Introducing Social Psychology. London: Penguin pp 427-433.

Tanner, C.A. (1977) The Effect of Hypothesis Generation as an Instructional Method on the Diagnostic Processes of Senior Baccalaureate Nursing Students. Unpublished Doctoral Dissertation. University of Colorado.

Taylor, S.E. & Crocker, J. (1981) Schematic Bases of Social Information Processing. R.T. Higgins, C.P. Herman & M.P. Zanna (Eds) Social Cognition: The Ontario Symposium, Vol I. Hillsdale N.J., Erlbaum.

Thiel, H.G., Parker, D. & Bruce, T.A. (1973) Stress Factors and the Risk of Myocardial Infarction. Journal Psychosomatic Research, 17, pp 43-57.

Truax, C.D. (1966) Inluence of Patient Statements and Therapist Statements During Psychotherapy. Journal Clinical Psychology, 22, pp 335-337.

Tulving, E. (1972) Episodic and Semantic Memory. E. Tulving & W. Donaldson (Eds). Organization of Memory. New York, Academic Press.

Tversky, A. & Kahneman, D. (1971) Belief in the Law of Small Numbers. Psychological Bulletin, 76, pp 105-110.

Tversky, A. & Kahneman, D. (1973) Availability: A Heuristic for Juding Frequency and Probability. Cognitive Psychology, 5, pp 207-232.

Tversky, A. & Kahneman, D. (1981) The Framing of Decisions and the Psychology of Choice. Science, 211, pp 453-458.

Van de Creek, L. & Watkins, J.T. (1972) Responses to Incongruent Verbal and Non-Verbal Emotional Cues. Journal of Communication, 22, pp 311-316.

Wallstan, T.S. (1972) Conjoint-Measurement Framework for the Study of Probabilistic Information Processing. Psychological Review, 79, pp 245-260.

Wallstan, T.S. (1977) Processing Information for Decisions. N.J. Castellan, D.B. Pisoni & G.R. Potts (Eds). Cognitive Theory, Vol 2. Hillsdale N.J. Erlbaum. ✓

Wallston, K.A., Wallston, B.S. & de Vellis, B.E. (1976) Effect of Negative Stereotype on Nurses Attitudes Towards an Alcoholic Patient. Journal Studies on Alcohol, 37, pp 659-665.

Ward, W.C. & Jenkins, H.M. (1965) The Display of Information and the Judgement of Contingency. Canadian Journal Psychology, 19, pp 231-241.

Wason, P.C. (1960) On the Failure to Eliminate Hypotheses in a Conceptual Task. Quart. J. Exper. Psychol. 12, pp 129-140.

Wason, P.C. (1966) Reasoning. B.M. Foss (Ed) New Horizons in Psychology. Harmondsworth, England. Penguin.

Wason, P.C. (1968) Reasoning About a Rule. Quart. J. Exper. Psychol. 20, pp 273-281.

Wason, P.C. (1969) Regression in Reasoning. British Journal Psychology, 60, pp 471-480.

Weinman, J. (1981) An Outline of Psychology as Applied to Medicine Bristol, John Wright.

Wells, T. (1975) Towards Understanding Nurses Problems in the Care of the Hospital Elderly. Unpublished PhD Thesis, University of Manchester.

Wilson-Barnett, J. (1977) Patients Emotional Reactions to Hospitalisation. Unpublished PhD Thesis, University of London.

Wilson-Barnett, J. & Carrigy, A. (1978) Factors Influencing Patients Emotional Reactions to Hospitalization. Journal of Advanced Nursing, 3, pp 221-229.

Wilson-Barnett, J. & Fordham, M. (1982) Recovery From Illness. Chichester. John Wiley & Sons.

Winer, B.J. (1962) Statistical Principles in Experimental Design. New York, McGraw-Hill Book Company.

Winkelwayer, R., Exline, R.V., Gottheil, E. & Paredes, A. (1978) The Relative Accuracy of U.S, British and Mexican in Judging the Emotional Displays of Schizophrenic and Normal U.S. Women. Journal Clinical Psychology, 34, pp 600-6008.

Wise, J.A. & Mockovak, W.P. (1973) Descriptive Modelling of Subjective Probabilities. Organizational Behaviour and Human Performance. 9, pp 292-306.

Zaidel, S. & Mehrabian, A. (1969) The Ability to Communicate and Infer Positive and Negative Attitudes Facially and Verbally. Journal Experimental Research in Personality, 3, pp 233-241.

Zung, W.W.K. (1973) From Art to Science. The Diagnosis and Treatment of Depression. Archives General Psychiatry, 29, pp 328-337.